

# 생성적 적대 신경망(GAN) 머신러닝을 활용한 글의 이미지화

2020년 전기 졸업과제 최종 발표  
(2020.09.25)

GAN다 박창조

정보컴퓨터공학전공 201524461 박 성국

정보컴퓨터공학전공 201724557 장 수현

정보컴퓨터공학전공 201724480 박 창조

# INDEX

1. 개요

2. 연구 사례

3. 설계

4. 구현

5. 분석 및 평가

6. 발전 방향

# 1. 개요

## 주제

생성적 적대 신경망(GAN) 머신러닝을 활용한 글의 이미지화

생성자  
(Generator)



GAN

Generative Adversarial Network

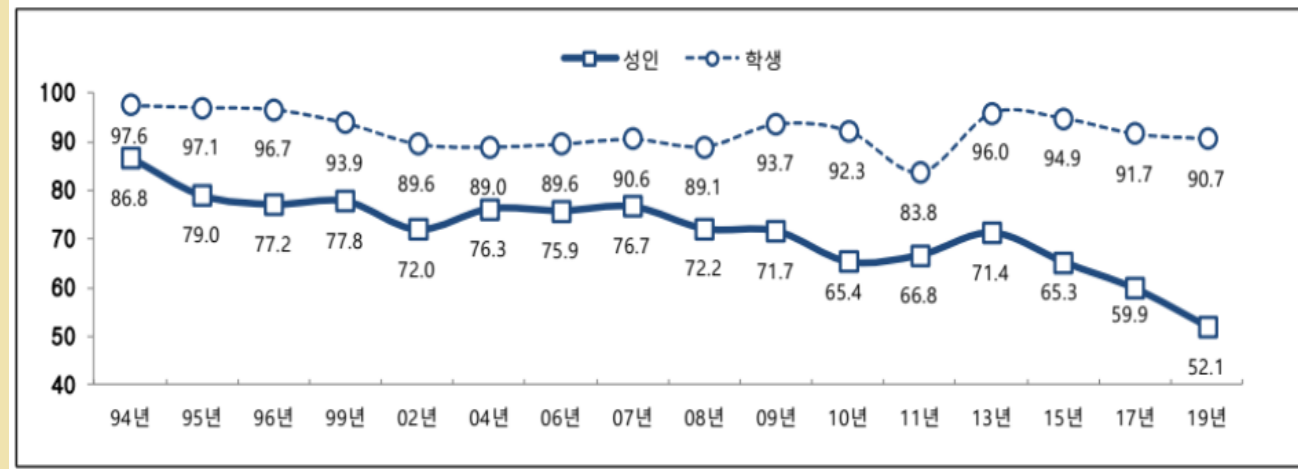


식별자  
(Discriminator)



# 1. 개요

## 배경



2019 국민 독서실태 조사에 따른 '종이책 독서율' 변화 추이(성인·학생)

01.24, 15:02  
누가 3줄 요약 좀 너무길어  
원문제목 : 고시생에게 버림받고 인생 파탄난 여성, 왕코너 [28]

---

01.23, 14:38  
이렇게 긴거 이해 못해요~ 3줄 요약 이거만 보거든요 ㅋㅋ... ㅋ ㅋ  
원문제목 : [꿀]문재인 체제개편안 관련 트윗 과연 누가 이를 비난할수 있단말인가 [46]

---

11.11, 16:37  
3줄 요약과 가려진 사진이 필요하다. ✨  
원문제목 : 현재 논란이 된 데뷔 예정 걸그룹 멤버 서자수 씨에 대한 피해자 측의 입장 정리입니다. (+음성 및 다른 증거. [13]

커뮤니티 '3줄 요약' 댓글 캡처 화면

12.17, 04:22  
글 넘 길어서 패스  
원문제목 : "그래 나 좋아한다 어필래?" [34]

---

12.02, 16:09  
요약머쥬? 저는 링크눌렀다가 너무길어서 패스했어  
원문제목 : 대략 40분 전에 자계에 어떤분이 [16]

---

11.10, 16:50  
너무길어서 패스  
원문제목 : 채식주의의 합정.bgm [65]

커뮤니티 '길어서 패스' 댓글 캡처 화면

# 1. 개요

## 목표



입력 - 객체 묘사 문장



출력 1 - 이미지 생성



출력 2 - 이미지 화풍 적용

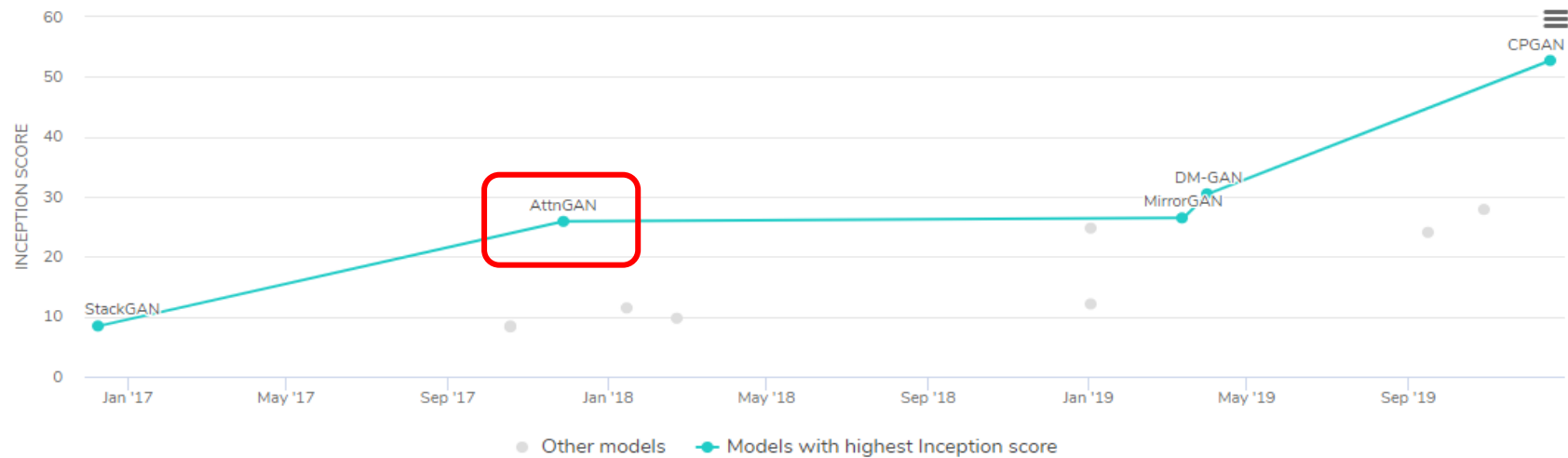
숲 속 어느 마을에 귀가 크고 뽕족  
하고, 눈과 코가 까맣고, 이마와  
꼬리는 갈색 털을, 배는 하얀 털을  
가진 여우가 살고 있었어요.



## 2. 연구 사례

### - 이미지 생성

#### Text-to-Image Generation on COCO



View

Inception score

All models

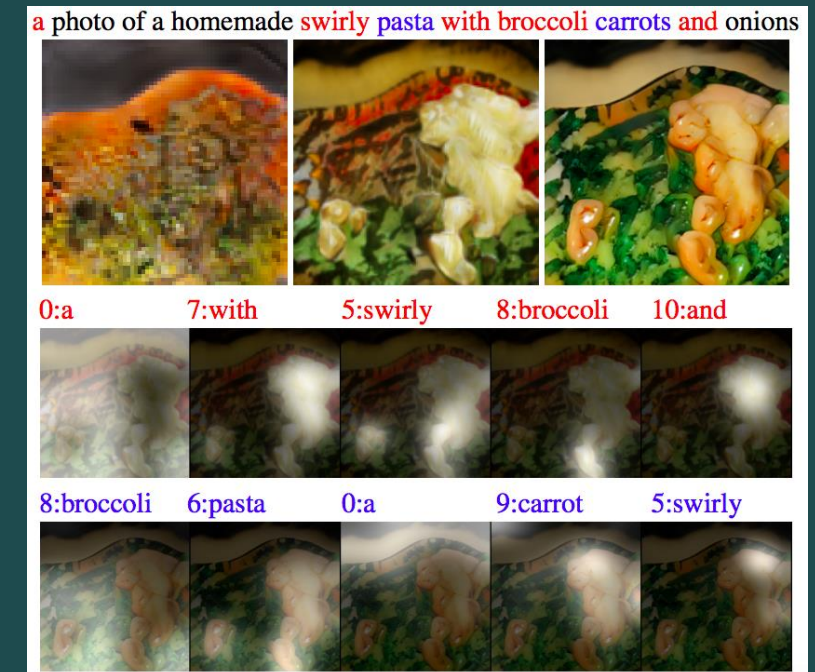
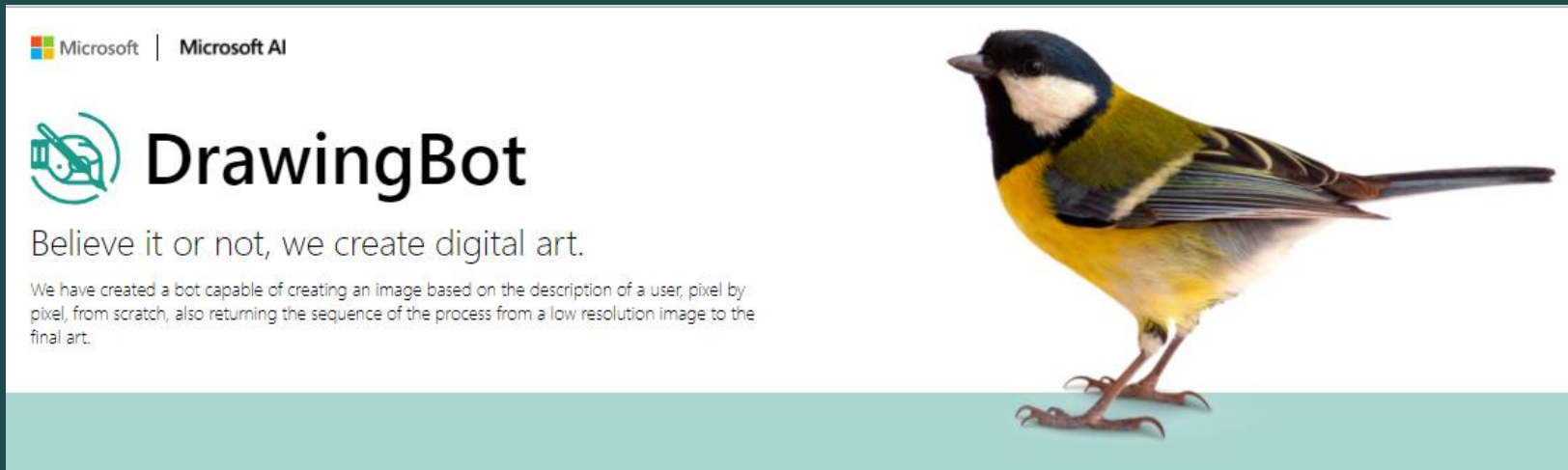
Edit

RANK	MODEL	FID	SOA-C	INCEPTION SCORE ↑	EXTRA TRAINING DATA	PAPER	CODE	RESULT	YEAR
1	CPGAN			52.73	×	CPGAN: Full-Spectrum Content-Parsing Generative Adversarial Networks for Text-to-Image Synthesis	<a href="#">GitHub</a>	<a href="#">Result</a>	2019
2	DM-GAN		33.44	30.49	×	DM-GAN: Dynamic Memory Generative Adversarial Networks for Text-to-Image Synthesis	<a href="#">GitHub</a>	<a href="#">Result</a>	2019
3	OP-GAN	24.70	35.85	27.88	×	Semantic Object Accuracy for Generative Text-to-Image Synthesis	<a href="#">GitHub</a>	<a href="#">Result</a>	2019
4	MirrorGAN			26.47	×	MirrorGAN: Learning Text-to-image Generation by Redescription	<a href="#">GitHub</a>	<a href="#">Result</a>	2019
5	AttnGAN		25.88	25.89	×	AttnGAN: Fine-Grained Text to Image Generation with Attentional Generative Adversarial Networks	<a href="#">GitHub</a>	<a href="#">Result</a>	2017

## 2. 연구 사례

### - 이미지 생성 ( Text-to-Image )

#### AttnGAN



- GitHub에 코드 제공
- 한글 사용 사례 無
- 인터페이스 미제공
- 화풍 미적용



## 2. 연구 사례

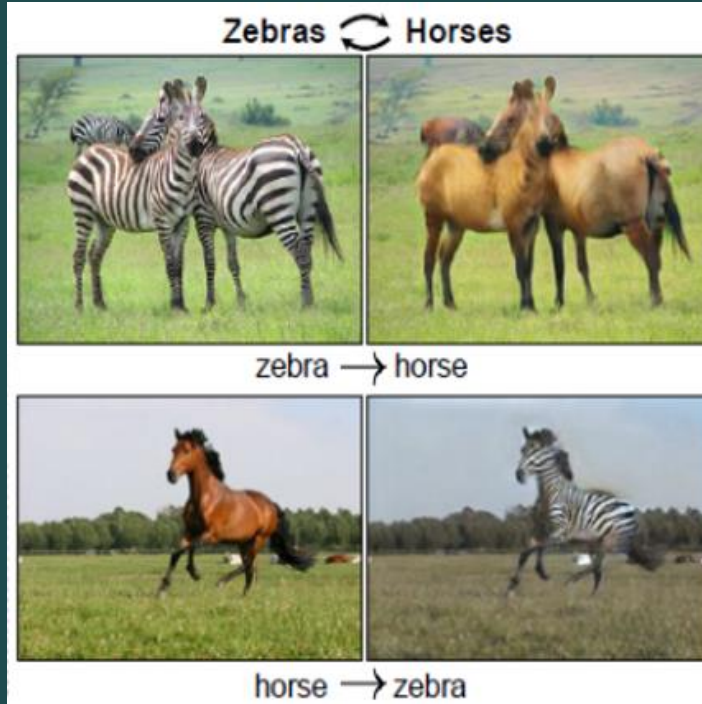
### - 화풍 적용 ( Image-to-Image )

pix2pix



- Paired Data

CycleGAN



- **Unpaired Data**
- GitHub에 코드 제공
- 인터페이스 미제공
- 이미지 생성 불가능



## 2. 연구 사례

### »» 시스템 기능 목표

AttnGAN, CycleGAN, 과제 시스템의 비교표

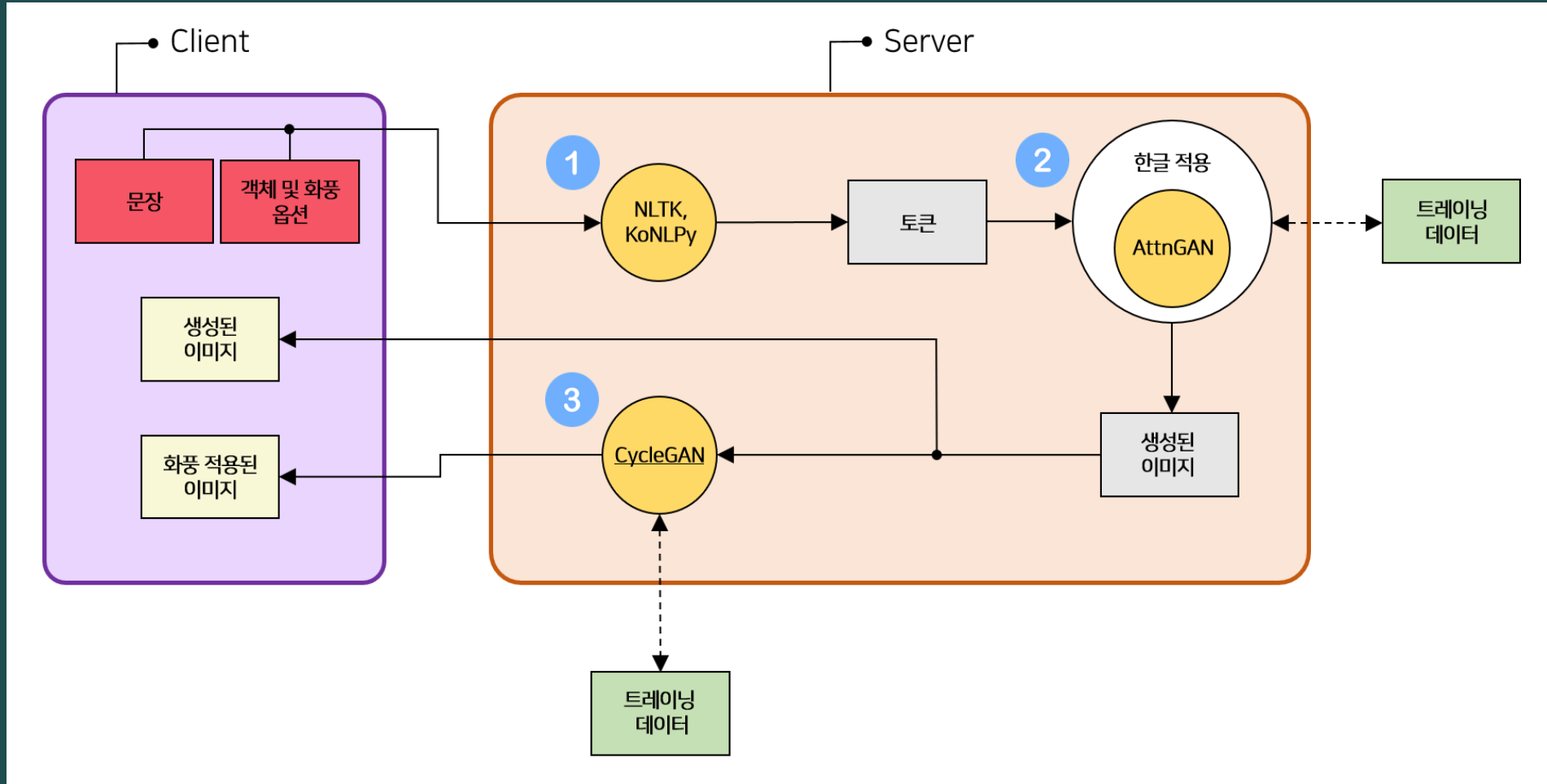
	AttnGAN	CycleGAN	과제 시스템
인터페이스 제공	X	X	O
한글 사용 가능	X	무관	O
이미지 생성	O	X	O
화풍 적용	X	O	O

- ☑ 사용자에게 웹 인터페이스 제공해야 한다.
- ☑ 사용 대상은 한국인으로 한글 입력이 가능하도록 해야 한다.
- ☑ 한 시스템 내에서 이미지 생성과 화풍 적용이 모두 이뤄져야 한다.

### 3. 설계

#### 전체 시스템 설계

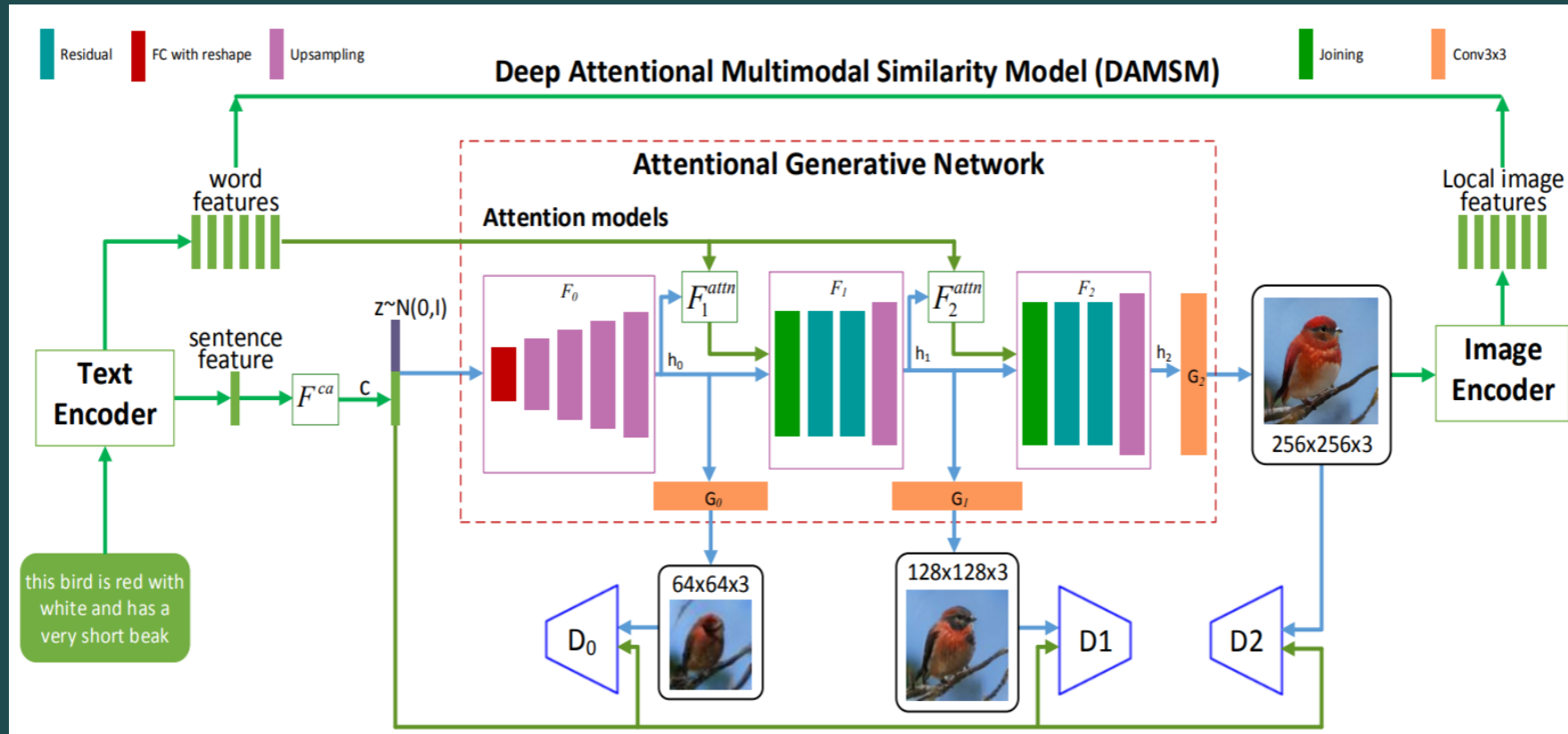
#### 시스템 DFD



### 3. 설계

#### 시스템 세부 설계 1

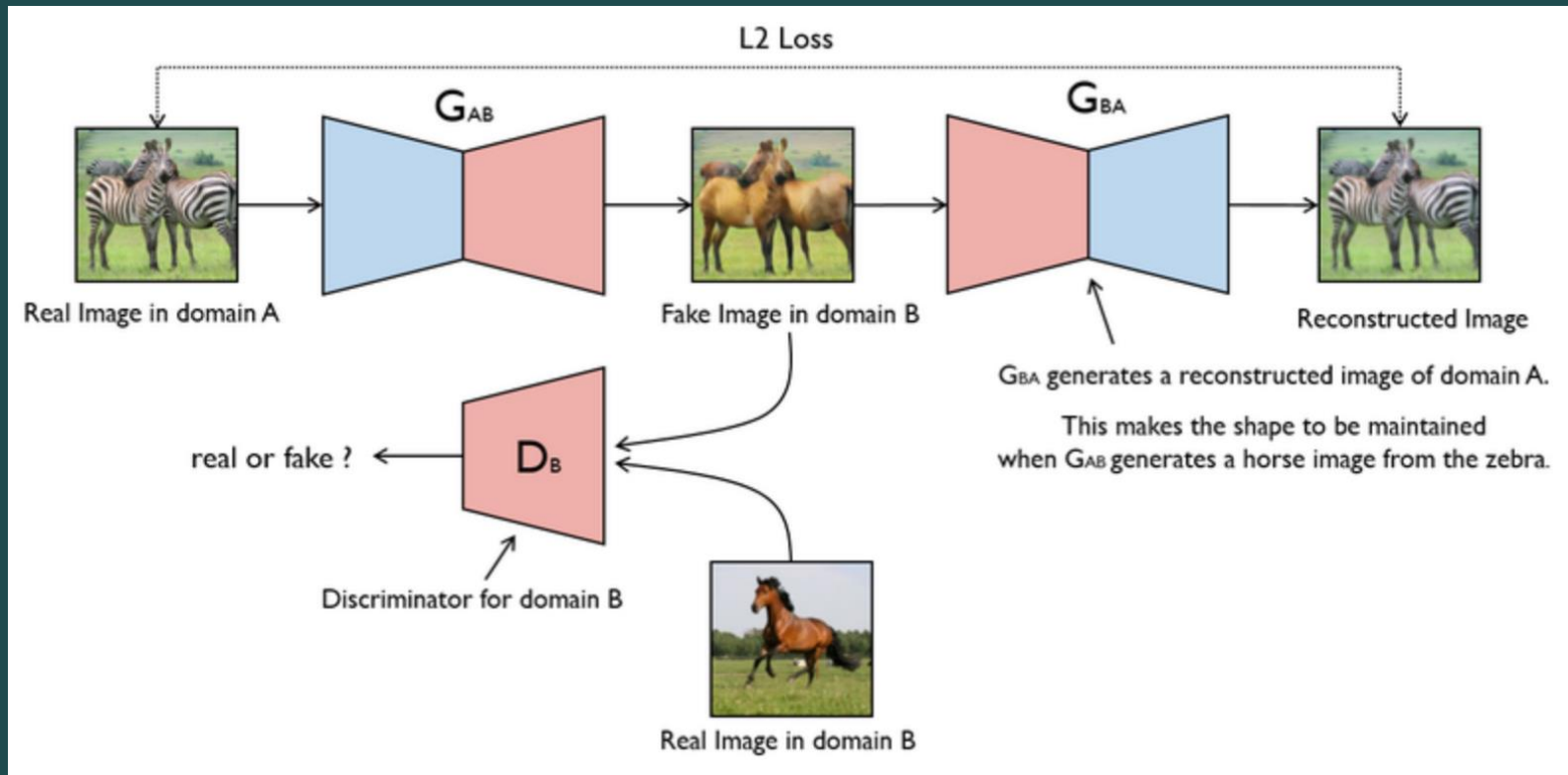
#### AttnGAN 구조



### 3. 설계

#### 시스템 세부 설계 2

#### CycleGAN 구조



## 4. 구현

### (1) 라이브러리 테스트 - 개발환경

※ 딥러닝에서는 CPU보다 GPU의 성능이 중요

- 구글 코랩

- 충분한 메모리

- 비교적 높은 GPU 성능

⇒ 트레이닝 가능

- 로컬

- 낮은 메모리

- 매우 낮은 GPU 성능

⇒ 사실상 트레이닝 불가능

colab



## 4. 구현

### (1) 라이브러리 테스트 - 초기 결과물

AttnGAN(초기)



이미지 2889장, 텍스트 28890문장

에포크: DAMSM-600, Attn-600

CycleGAN(초기)



사과 이미지 1109장, 오렌지 이미지 928장

에포크: 100

⇒ 낮은 퀄리티

## 4. 구현

### (2) CycleGAN 동화풍 데이터 수집



- 동화 풍 이미지
  - 육안으로 봤을 때 비슷한 느낌의 이미지 수집
- ⇒ 동화 풍 데이터 약 3000장 수집 후 품질이 떨어지는 데이터 약 2000장을 제외하여 양질의 이미지 약 1000장 선정



## 4. 구현

### (2) CycleGAN 동화풍 데이터 수집



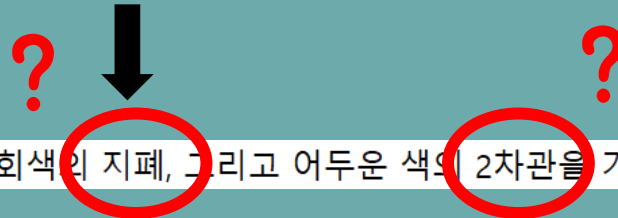
⇒ 하지만 **절대적인 트레이닝 횟수 부족**으로 인해 눈에 띄는 차이가 나타나지 않음.

## 4. 구현

### (3) AttnGAN 한글화 문제

텍스트 데이터 오 번역

the small bird has a large white eye, a short orange and gray bill, and dark colored secondaries.



- Bill(새의 부리)이 '지폐'로 번역
- Secondaries(보조)가 '2차관'으로 번역

부적합한 형태소 분석기



한글 자연어 처리에 부적합

## 4. 구현

### (3) AttnGAN 한글화 문제

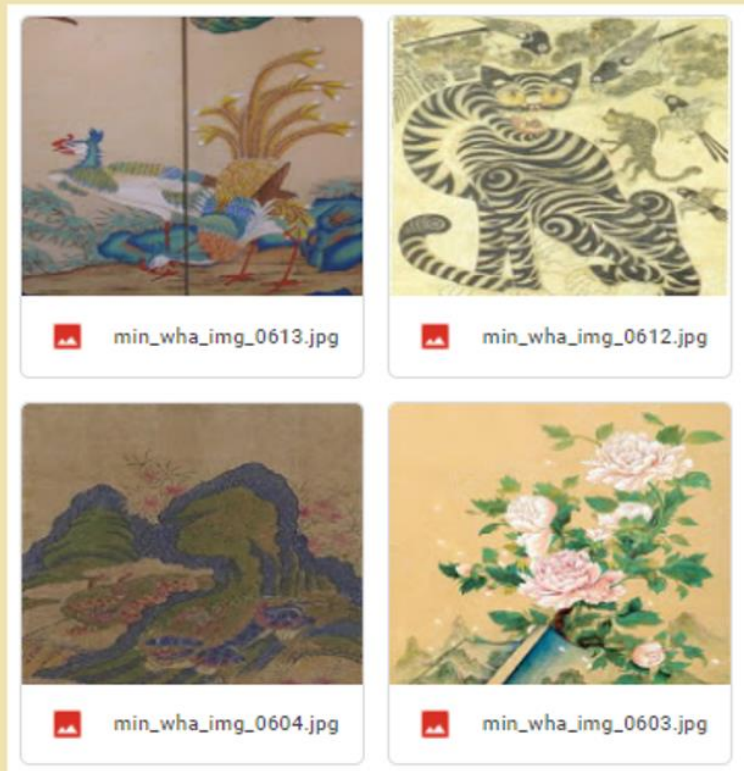
이미지 약 2800장, 한글 텍스트 약 28000문장 트레이닝,  
DAMSM: epoch 600, Attn: epoch 500



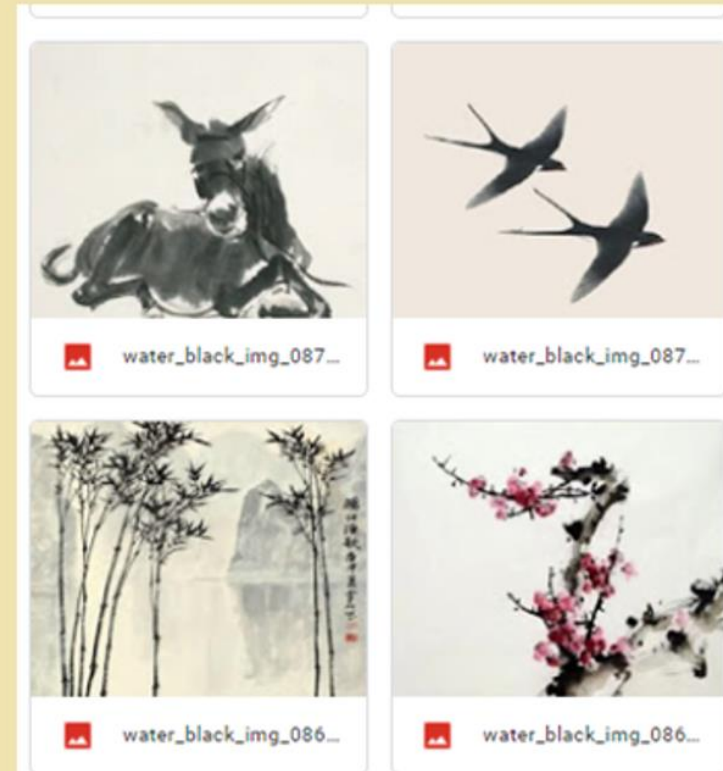
⇒ 새라고 보기에는 **너무 낮은** 퀄리티

## 4. 구현

### (4) 시도 1 : CycleGAN 화풍 다양화



동양화 데이터셋 일부



수묵화 데이터셋 일부

## 4. 구현

### (4) 시도 2 : AttnGAN 한글 데이터 품질 개선



- 파파고 번역 결과 **오역**이 다소 발견됨.
  - 같은 의미의 단어를 **다른 단어로 번역**
- ⇒ 구글 번역기를 통해 **재 번역**

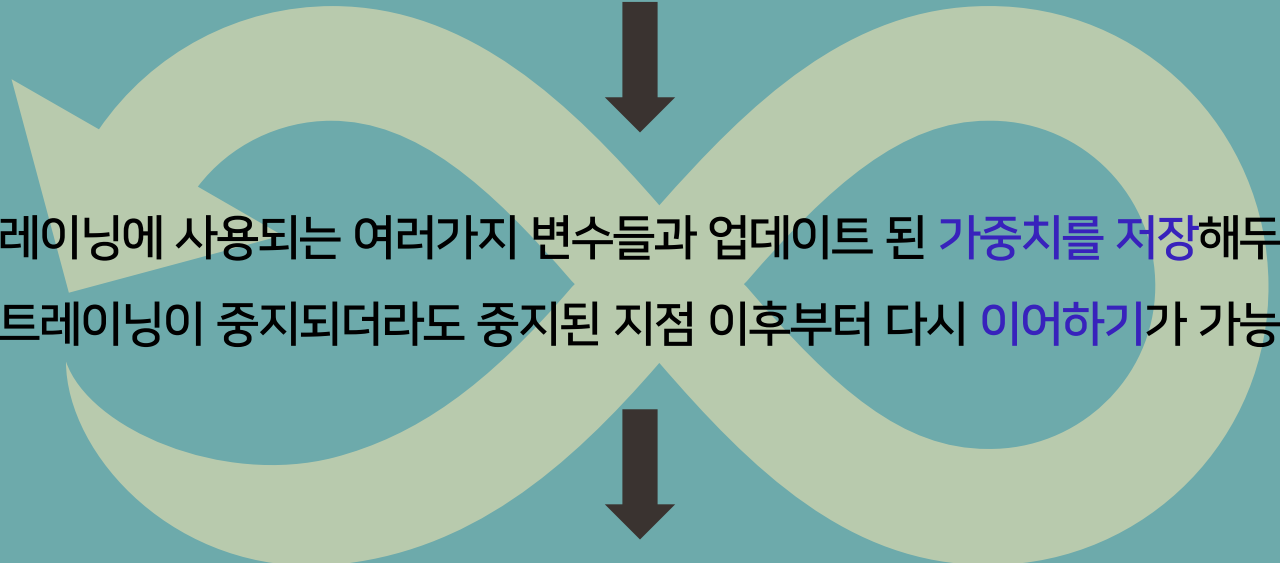


- 기존의 NLTK는 영어 문장 분석에 적합
  - **한글 문장 분석에 적합**한 KoNLPy
  - 명사, 형용사, 동사에 해당하는 품사만 포함
- ⇒ NLTK를 **KoNLPy**로 교체

## 4. 구현

### (4) 시도 3 : 트레이닝 이어서 하기

여러가지를 개선했지만 근본적으로 **절대적인 트레이닝 횟수 부족**



트레이닝에 사용되는 여러가지 변수들과 업데이트 된 **가중치를 저장**해두면  
트레이닝이 중지되더라도 중지된 지점 이후부터 다시 **이어하기**가 가능

트레이닝 횟수에 사실상 **제한이 없어지면서** 결과물의 **정확도와 퀄리티가 크게 향상**



## 4. 구현

### (5) 패키징 및 웹 호스팅



- Python으로 작성된 마이크로 웹 프레임워크
  - Werkzeug 툴킷과 Jinja2 템플릿 엔진에 기반
- ⇒ Python으로 작성된 프로그램과 웹을 연결



- Python을 기반으로 하는 웹 호스팅 서비스
  - 온라인 통합개발환경 및 Bash 인터페이스 제공
- ⇒ 작성한 프로그램을 웹으로 호스팅



## 4. 구현

### (6) 최종 구현 결과 - AttnGAN

#### AttnGAN

- 새(bird) 이미지 생성 가능(정확도 **비교적 높음**)
- 토끼 이미지 생성 가능(정확도 **비교적 낮음**)
- 호랑이 이미지 생성 가능(정확도 **매우 낮음**)

새

토끼

호랑이



## 4. 구현

### (6) 최종 구현 결과 - CycleGAN

원본

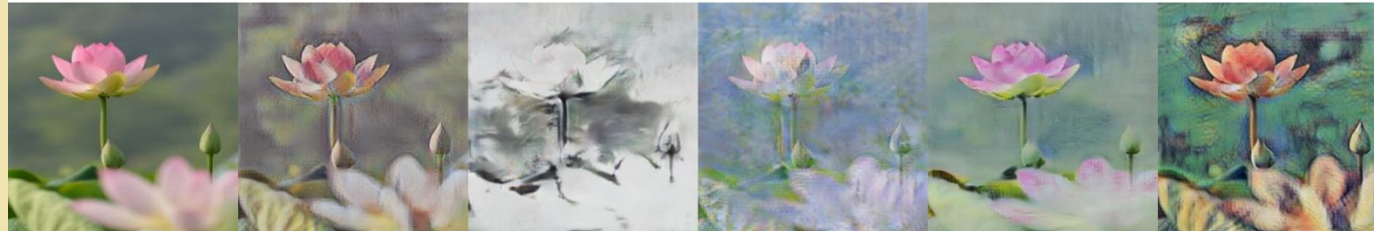
세잔

수묵화

모네

동양화

반고흐

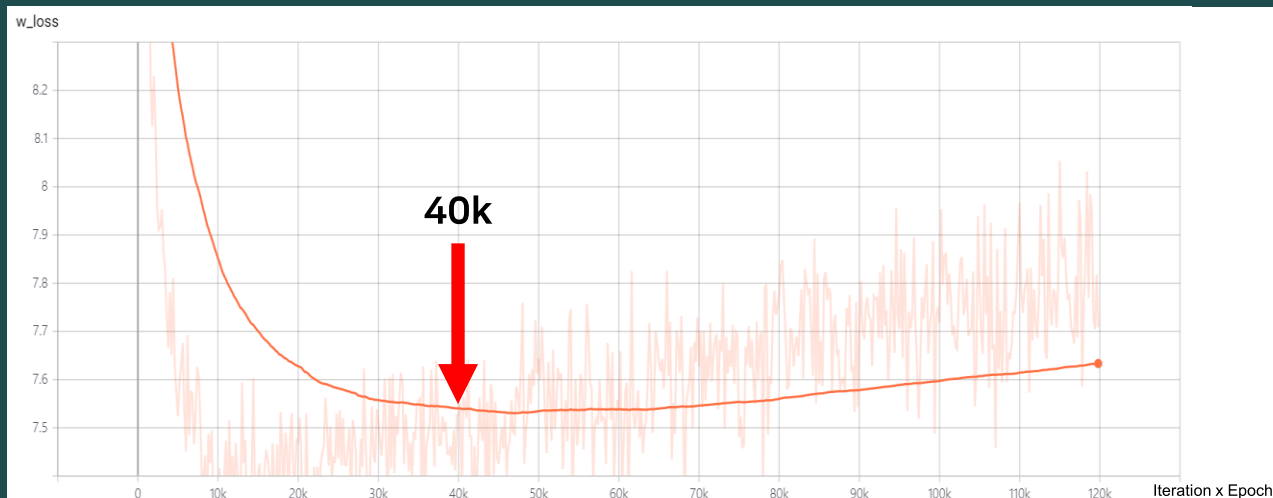
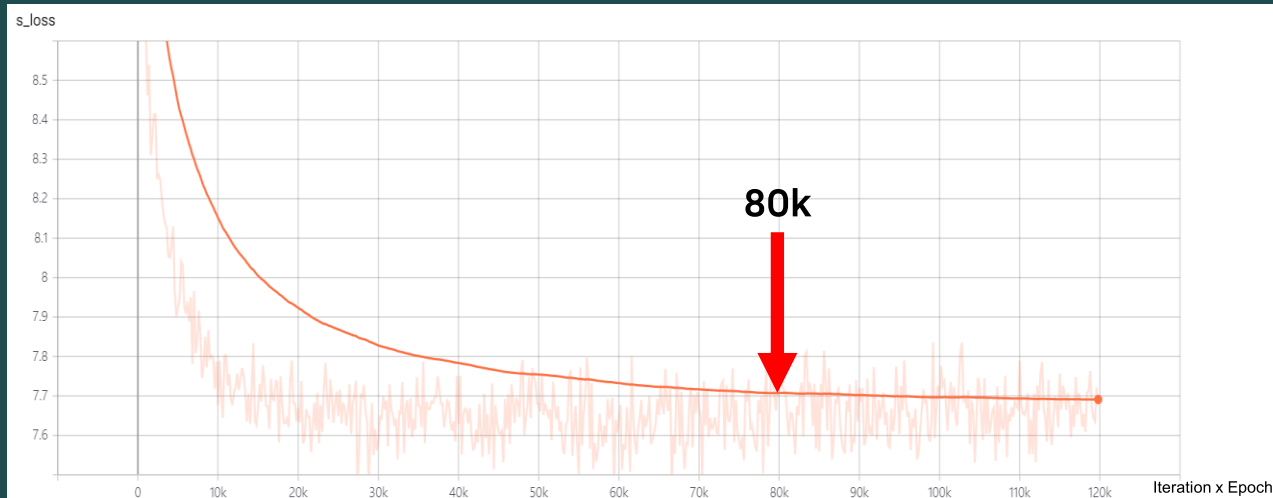


### CycleGAN

- 동양화, 수묵화, 세잔, 반 고흐, 모네 화풍 적용 가능
- 이미지에 따라 **변화의 정도에 차이**가 있음

## 5. 분석 및 평가

### AttnGAN - DAMSM



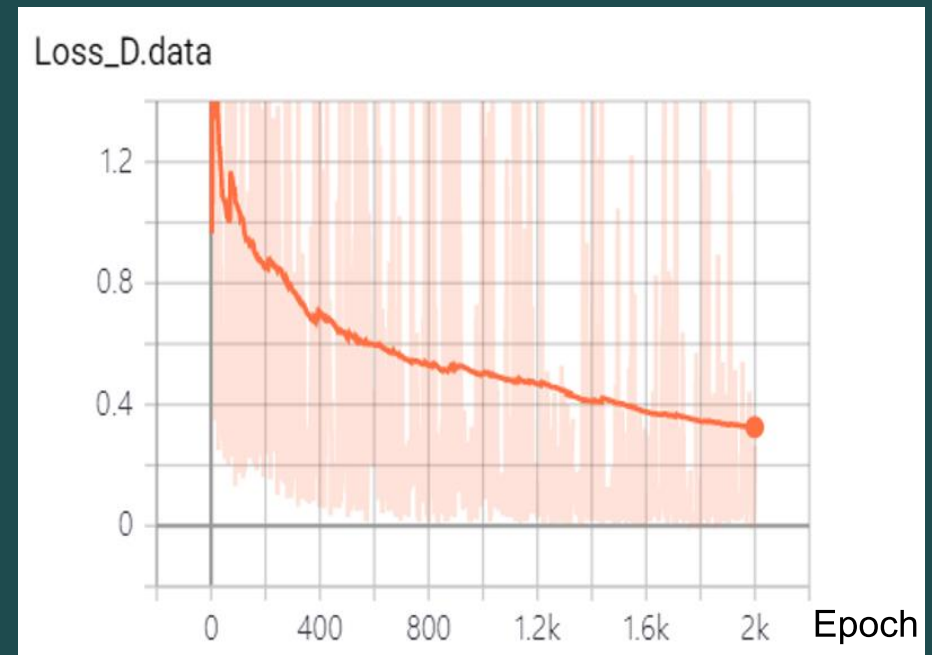
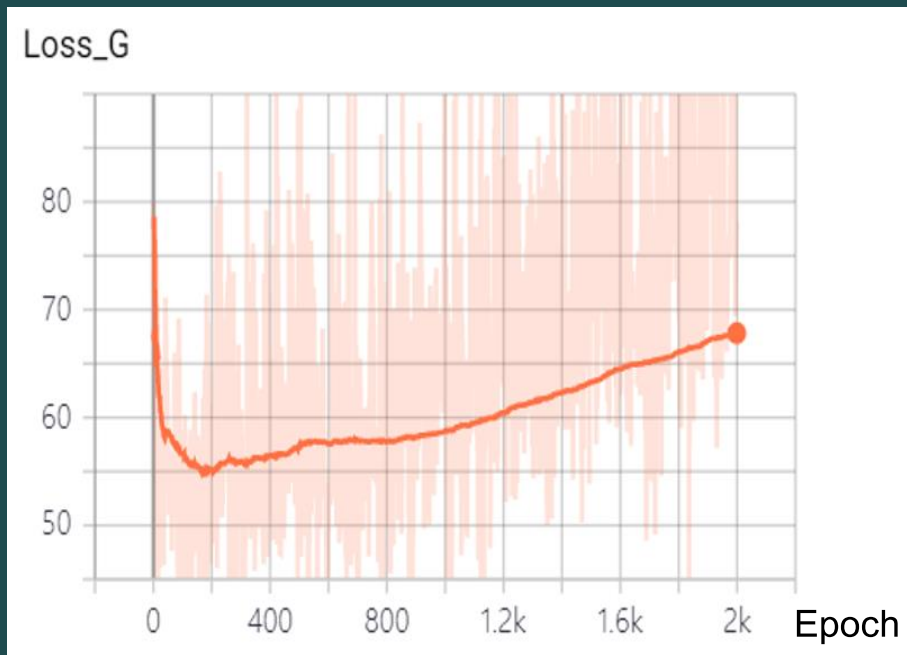
- 평균값인 60k를 최적값으로 설정
- Iteration x Epoch = 60000  
Iteration = 30  
Epoch = 2000
- 최적 epoch를 2000으로 설정

- S\_Loss : 문장 레벨 손실
- W\_Loss : 단어 레벨 손실

## 5. 분석 및 평가

### AttnGAN – Attentional Generative Network

- Loss\_G : 생성자(Generator)의 손실 함수(Loss Function) 값
- Loss\_D : 식별자(Discriminator)의 손실 함수(Loss Function) 값



Loss 값이 아직 수렴하지 않았기 때문에 현재 Epoch를 최적의 Epoch로 설정

## 5. 분석 및 평가

### AttnGAN 결과

- 새



- 토끼



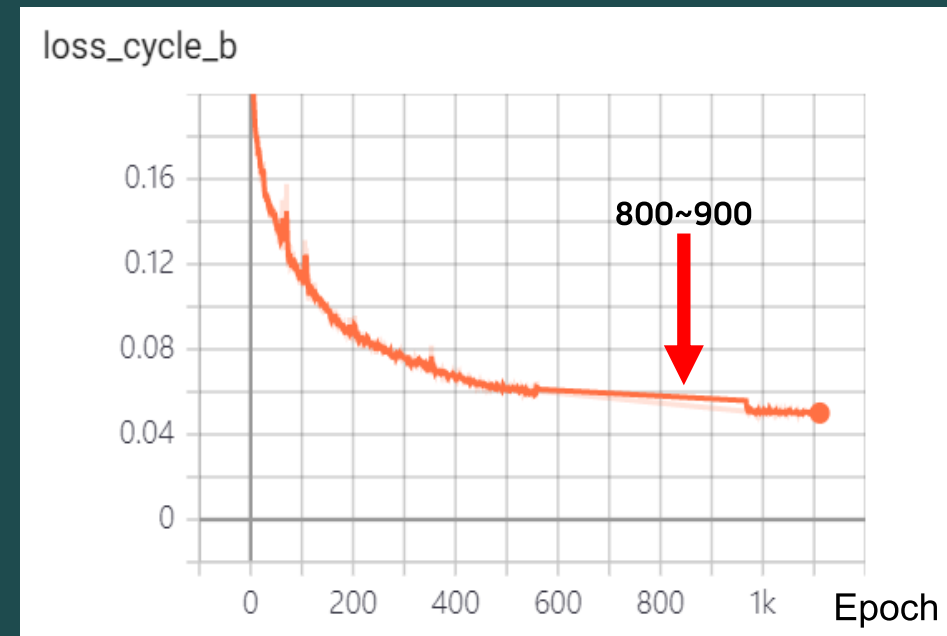
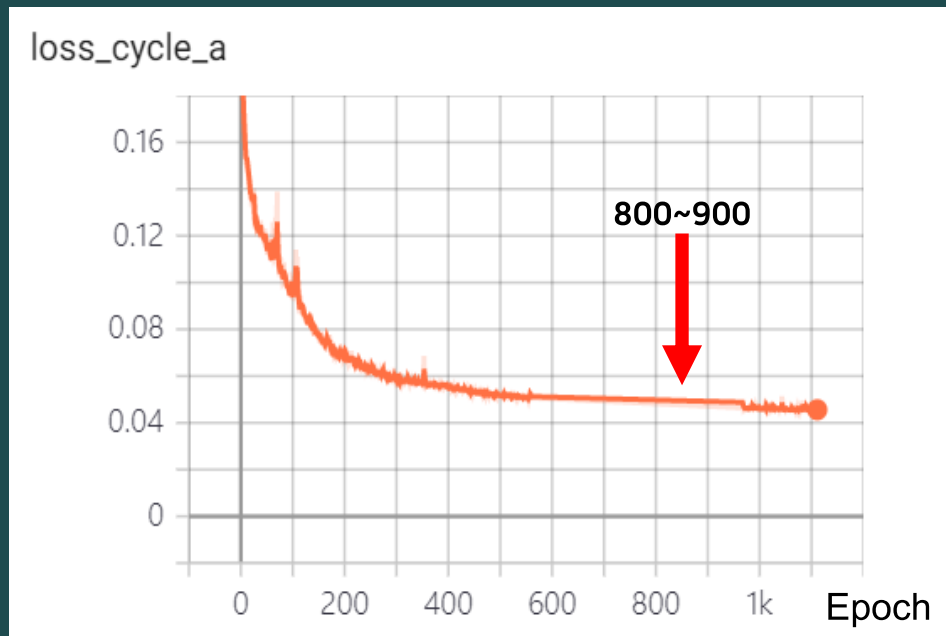
- 호랑이



## 5. 분석 및 평가

### CycleGAN

- `loss_cycle_a` : 이미지를 도메인 A 이미지를 도메인 B로 변환 후 다시 A로 변환했을 때의 손실 함수 값
- `loss_cycle_b` : 이미지를 도메인 B 이미지를 도메인 A로 변환 후 다시 B로 변환했을 때의 손실 함수 값



Loss 값이 수렴하기 시작하는 지점의 Epoch를 최적의 Epoch로 설정



## 5. 분석 및 평가

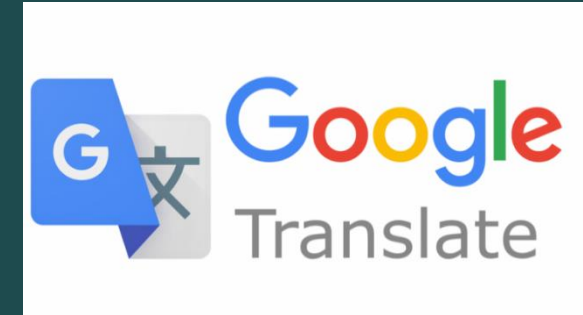
### 한글 토큰 정확도



#### <실험 환경>

입력 데이터: 한글 문장 약 3000개

\* 명사, 형용사, 동사만 포함



# VS

- 전체 토큰 개수 : 311,586개
- 5회 이하의 토큰 개수 : 4,805개
- 전체 토큰에 대한 비율 : **1.542%**

- 전체 토큰 개수 : 283,181개
- 5회 이하의 토큰 개수 : 3,361개
- 전체 토큰에 대한 비율 : **1.187%**



## 5. 분석 및 평가

### 평가

#### 전체적인 결과에 대한 평가

- 새에 대한 사진 품질은 양호
- 다른 동물에 대한 사진 품질은 비교적 낮음
- 화풍 적용 후 이미지 품질은 매우 만족스러운 결과
- 웹 디자인 또한 사용자 친화적으로 만족스러운 결과

#### 아쉬운 점

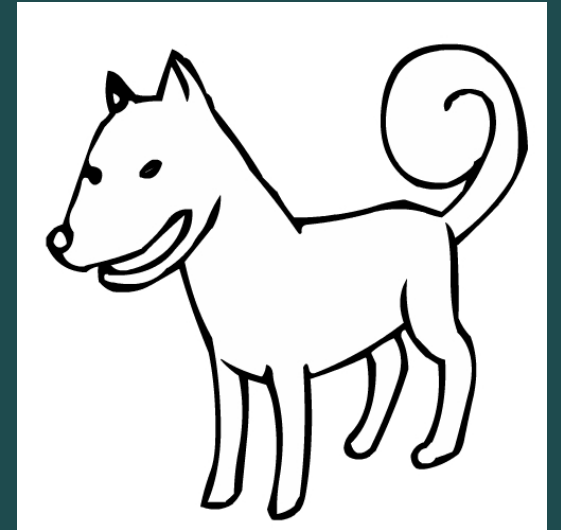
- 데이터 로딩
  - 다른 방법을 모색하는데 많은 시간 소비
- 데이터 품질
  - 정렬된(Aligned) 이미지 데이터 수집
  - 일관성 있는 텍스트

## 6. 발전 방향

### 향후 과제



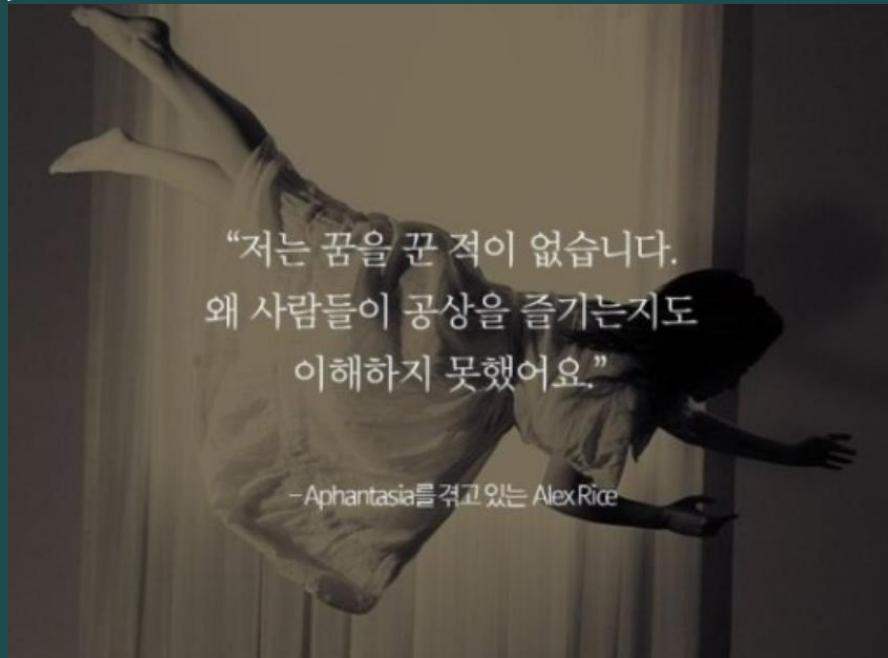
객체 중심이 아닌 술어 중심의 변환



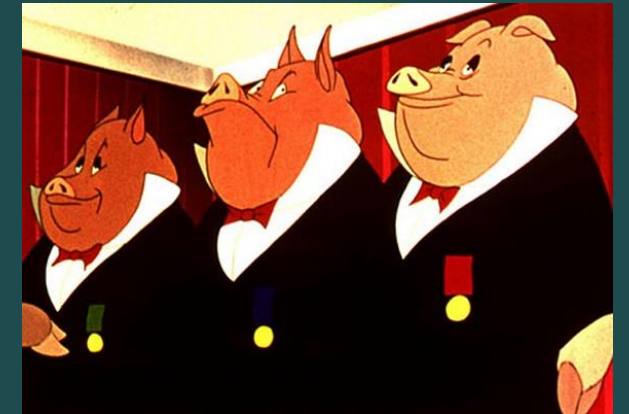
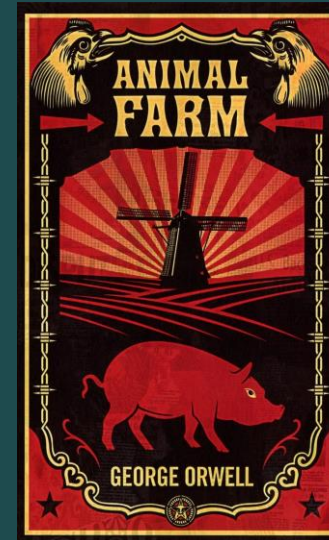
단순 색 변환 뿐 아니라 윤곽선, 형태 등의 변환

## 6. 발전 방향

### 활용 방안



난독증, 아판타시아 병을 앓는 환자들을 위한 도움



글만 있는 소설이나 책에 자동으로 그림 삽입

# 역할 분담

## 박 성국

## 장 수현

## 박 창조

### • AttnGAN 한글화 작업

- 영어 텍스트 데이터 한글화
- AttnGAN 한글화
- 다른 객체 한글 텍스트 데이터 생성
- AttnGAN 동물 데이터 한글 학습
- Konlpy 라이브러리 확인 및 변경

### • AttnGAN 이미지 생성 작업

- AttnGAN 이미지 품질 향상

### • CycleGAN 화풍 변환 작업

- CycleGAN 라이브러리 테스트
- 이미지 화풍 변환 테스트
- 화풍 이미지 데이터 수집
- CycleGAN 화풍 데이터 학습
- 세잔, 모네, 고흐 화풍 적용

### • AttnGAN 한글화 작업

- 다른 객체 한글 텍스트 데이터 생성

### • 웹 UI 생성 및 모듈 통합 작업

- 웹 인터페이스 생성
- 웹 인터페이스, 모듈 통합
- 서버 개설 및 웹 호스팅

### • AttnGAN 이미지 생성 작업

- 다른 객체 이미지 데이터 수집

### • CycleGAN 화풍 변환 작업

- 화풍 이미지 데이터 수집
- 객체 이미지 데이터 수집
- CycleGAN 화풍 데이터 학습
- 수묵화, 동양화 화풍 적용
- CycleGAN 화풍 품질 향상

### • AttnGAN 한글화 작업

- 다른 객체 한글 텍스트 데이터 생성
- AttnGAN 객체 데이터 한글 학습

### • 환경 개선을 위한 작업

- Google Cloud Platform 환경 확인

### • AttnGAN 이미지 생성 작업

- AttnGAN 라이브러리 테스트
- AttnGAN 최소 데이터 set 측정
- 기존 동물 이미지 생성 테스트
- 다른 객체 이미지 데이터 수집
- AttnGAN 객체 데이터 영어 학습
- AttnGAN 이미지 생성 품질 향상



감사합니다.

GAN다 박창조

정보컴퓨터공학전공 201524461 박 성국

정보컴퓨터공학전공 201724557 장 수현

정보컴퓨터공학전공 201724480 박 창조