딥러닝 기반 스테가노그래피 및 검출 시스템

팀명: 삼인조, 팀원: 안지수, 고가은, 곽지현

INDEX

01

연구 배경

- 최근 사이버 스파이 조직이 스테가노그래피 기술을 활용하여 정보를 은밀히 전달

- EXIF 메타데이터, 이미지 내부에 악성코드 삽입 사례 다수 발견

- 암호화+은닉 기법 결합으로 탐지 회피 기술 진화

- 사이버 공격자들 사이에서 스테가노그래피 기법이 확산 중

01

주제 선정 이유

안전성

의료 데이터 보호, 디지털 저작권,
군사 통신 등 다양한 분야에 활용

악의적 사용 증가

악성코드 은닉이나 랜섬웨어 유포 등에
악용되는 사례 증가

딥러닝 기술 발전

기존보다 더 정교하게 생성 및 탐지
정확도가 매우 높아짐

01

용어 정의 (1)

스테가노그래피 (Steganography)

메시지나 데이터를 이미지, 오디오, 비디오 파일 등의 다른 파일 안에 감추어 두는 기술

스테그아날리시스 (Stegalysis)

스테가노그래피로 숨겨진 데이터를 감지하고 복구하기 위한 기술

01

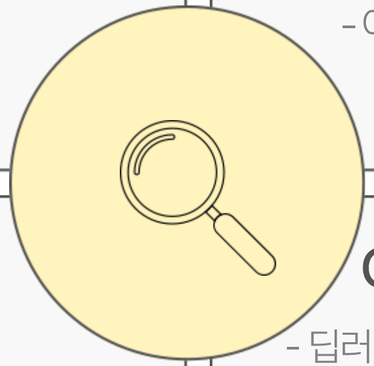
용어 정의 (2)

GAN

- 서로 대립하는 두 시스템(생성기, 판별기)의 경쟁을 통해 학습하는 방법
- 비지도학습으로 입력 값만으로 학습이 가능
- 지도학습의 초기 조건에 대한 민감도가 높음

생성기(Generator)

- 이미 존재하는 데이터와 비슷한 모조 데이터를 생성



판별기(Discriminator)

- 입력 데이터가 실제 데이터인지 구별하는 역할을 수행하고, 50% 확률로 진위 여부 판단하면 학습을 종료

CNN

- 딥러닝에서 주로 이미지나 영상 데이터를 처리할 때 사용 합성곱 계층과 풀링 계층을 통해 특징을 자동으로 추출하고, 완전 연결 계층을 통해 분류를 수행한다.

02

프로젝트 목표

- GAN기반 자연스러운 Stego 이미지 생성
- CNN기반 Stego 탐지 모델 구현
- 통합 프로그램 구현

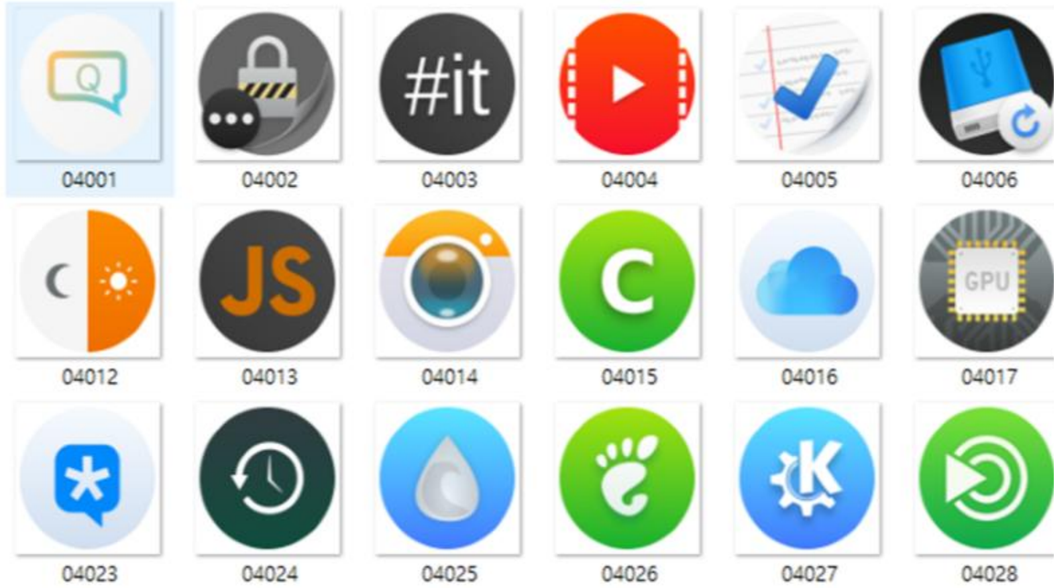
03

시스템 구조

- 데이터셋
- GAN 구조
- CNN 구조

03

데이터셋



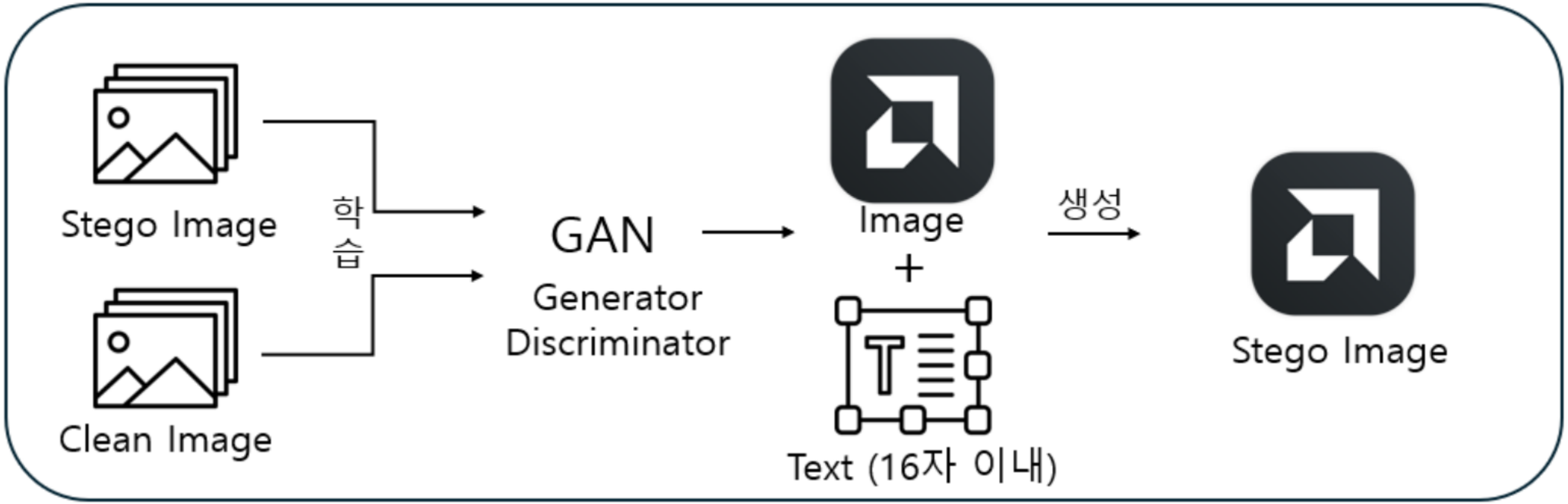
데이터셋(dataset)

=> Clean 8,000장 + Stego 6,000장 (512×512)

총 14,000장 이미지 사용 (출처: Kaggle)

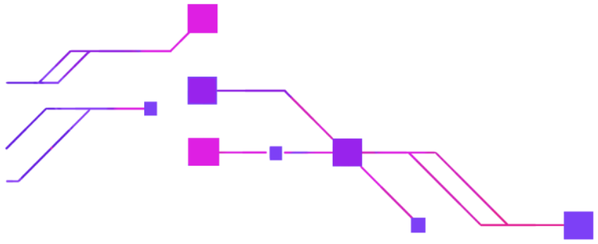
03

시스템 구조 (GAN)


Stego Image
Clean Image
학습
GAN
Generator
Discriminator
Image
+
Text (16자 이내)
생성
Stego Image

03

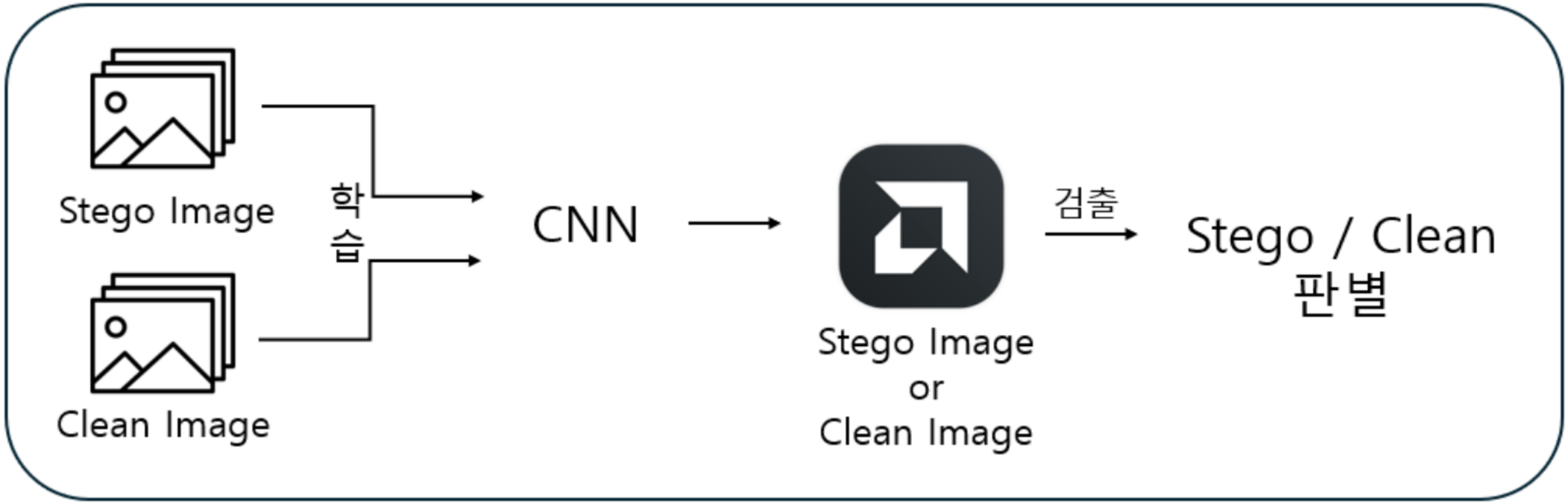
GAN 모델



01	U-Net 기반 Generator + CBAM
02	입력 : 512X512 이미지 + 16자 텍스트
03	출력 : 시각적으로 유사한 stego 이미지
04	손실함수 : Adversarial / Reconstruction / Perceptual

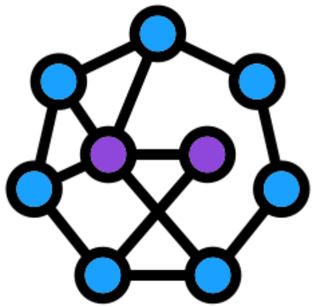
03

시스템 구조 (CNN)


Stego Image
Clean Image
학습
CNN
Stego Image
or
Clean Image
검출
Stego / Clean
판별

03

CNN 모델



01	ResNet + CBAM
02	입력 : clean 이미지 or stego 이미지
03	출력 : stego 이미지 판별
04	손실함수 : FocalLoss + 가중치 적용

04

실험 결과 및 성능 평가

04

실험 결과

GAN

G Loss : 0.0669

D Loss : 0.7214

PSNR : 31.81

SSIM : 0.9760

CNN

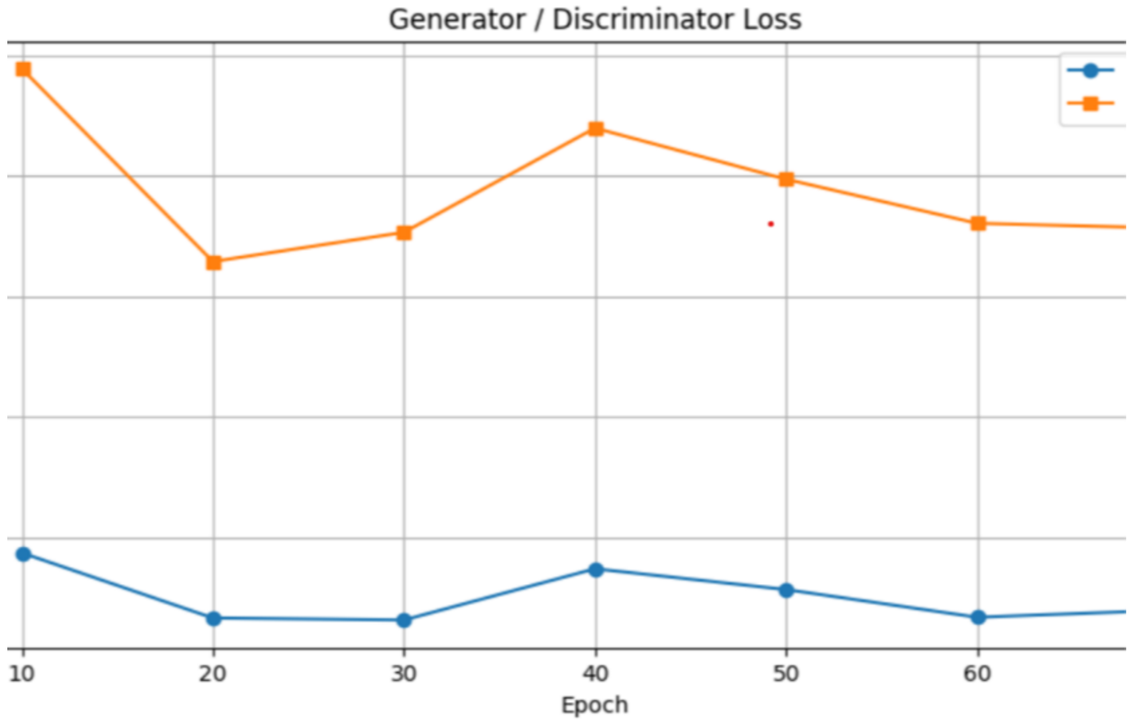
Accuracy : 0.8381

AUC : 0.9299

Loss : 0.28

04

성능 평가



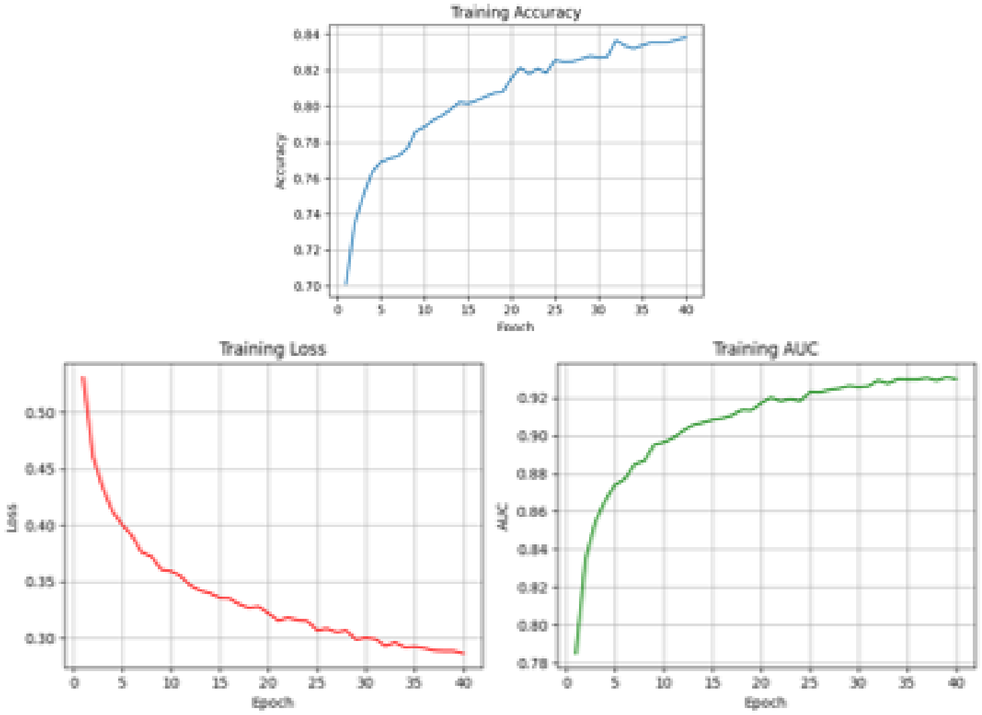
GAN 그래프

G Loss

D Loss

04

성능 평가



CNN 그래프

Accuracy

Loss

AUC

05

프로그램 개발

05

프로그램 개발

GAN Steganography

생성 전 이미지

생성 후 이미지

이미지 선택 후
숨기려는 텍스트 (영어 16자 이내) 입력

=> Stego 이미지 생성

05

프로그램 개발

CNN Steganalysis

Stego 의심되는 이미지 선택 후
결과 보기 버튼을 통해 확인 가능

06

시연 영상 및 기대효과

06

시연 영상


통합 스테가노그래피 시스템
텍스트 은닉
탐지
GAN Steganography
이미지 선택
파일 경로 없음
찾기
텍스트 (16자 이내)
Stego 이미지 생성
[결과 대기 중]

06

기대 효과

01	보안성 강화
02	다양한 산업 분야 응용
03	탐지 기술 고도화
04	차세대 보안 기술로의 가능성

감사합니다