

Mu-seong's Portfolio

꾸준히 꿈을 향해

목차

1. Research
 - User Experience Development Process for User Needs-based Chatbot
 - Natural Language Processing Text Augmentation Using Hierarchy-based Word Replacement
 - Multi-Label Classification Label Embedding Using AutoEncoder with Skip-Connections
 - Computer Vision Two-Stage Image Augmentation Methodology Using Style Transfer
2. Project
 - 항공 빅데이터 및 다층 신경망 모형을 활용한 실시간 3D 운항분석 및 예측 플랫폼 개발
 - 에스테틱 관리를 위한 진단 및 관리 후 사진 생성 모듈 구현
 - 지능형 사회안전 시스템 구축을 위한 멀티모달 딥러닝 기반 심층 상황 이해 모델 개발

1. User Experience

Development Process for User Needs-based Chatbot: Focusing on Design Thinking Methodology

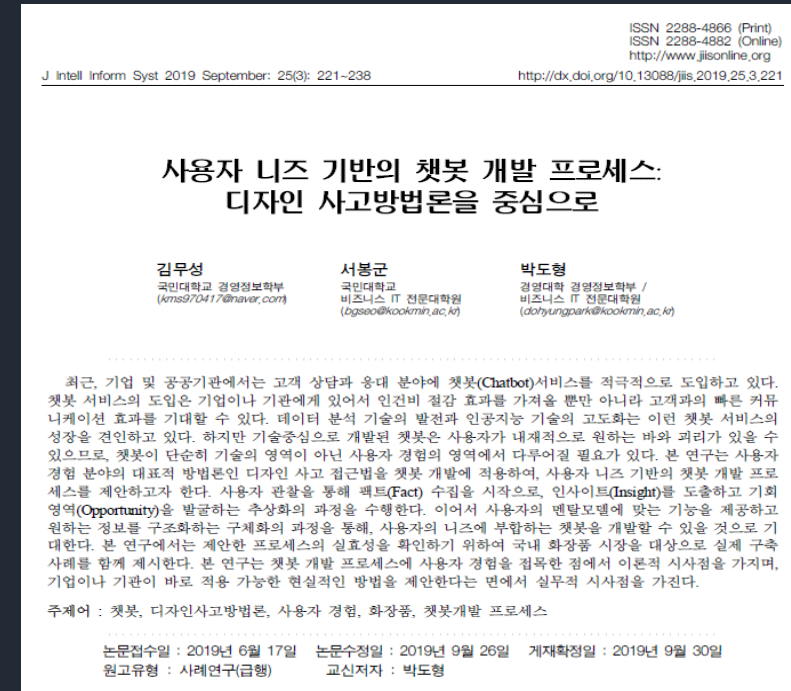
- 연구 기간 : 2018.03. ~ 2019.05.
- 역할 : 주저자
- 프레임워크 : Google DialogFlow
- 2018.03. ~ 2018.09. 화장품 제조업체 협업 프로젝트 수행
- 2019.06. 한국지능정보시스템학회 춘계학술대회 발표
- 2019.09. Journal of Intelligence and Information Systems (KCI, Publication)

• 연구 문제

- ✓ 기존 챗봇은 기술 중심으로 개발되는 경우가 많음
- ✓ 사용자와 밀접한 관련이 있음에도 불구하고 사용자 경험이 반영된 챗봇 연구는 미비함

• 연구 제안 내용

- ✓ 디자인 사고 접근법을 챗봇 개발에 적용함
- ✓ 사용자 니즈 기반의 챗봇 개발 프로세스 제안 및 챗봇 을 개발함



- 연구 기간 : 2020.03. ~ 2020.10.
- 역할 : 주저자
- 사용 언어 및 프레임워크 : Python, Pytorch
- 2020.11. 한국지능정보시스템학회 춘계학술대회 발표
- 2021.01. Journal of The Korea Society of Computer and Information (KCI, Publication)
- 2021.07. 계층 기반 단어 대체를 통한 텍스트 데이터 증강 장치 및 방법 특허 출원

• 연구 문제

- ✓ Text to Image 합성은 방대한 양의 이미지와 텍스트 데이터가 학습에 필요함
- ✓ 기존의 텍스트 증강 방법은 증강 과정에서 의미가 왜곡될 가능성이 높음

• 연구 제안 내용

- ✓ 단어가 갖는 품사별 특징을 활용하여 증강된 텍스트 데이터의 의미를 최대한 보존함
- ✓ 명사는 상위어로, 이외의 품사는 유의어로 대체하는 의미 계층 기반의 텍스트 증강을 제안함

JKSCI

한국컴퓨터정보학회논문지
Journal of The Korea Society of Computer and Information
Vol. 26 No. 1, pp. 0-0, January 2021
<https://doi.org/10.9708/jksci.2021.26.01.000>

Text Augmentation Using Hierarchy-based Word Replacement

Museong Kim*, Namgyu Kim*

*Graduate Student, Graduate School of Business IT, Kookmin University, Seoul, Korea

*Professor, Graduate School of Business IT, Kookmin University, Seoul, Korea

[Abstract]

Recently, multi-modal deep learning techniques that combine heterogeneous data for deep learning analysis have been utilized a lot. In particular, studies on the synthesis of Text to Image that automatically generate images from text are being actively conducted. Deep learning for image synthesis requires a vast amount of data consisting of pairs of images and text describing the image. Therefore, various data augmentation techniques have been devised to generate a large amount of data from small data. A number of text augmentation techniques based on synonym replacement have been proposed so far. However, these techniques have a common limitation in that there is a possibility of generating a incorrect text from the content of an image when replacing the synonym for a noun word. In this study, we propose a text augmentation method to replace words using word hierarchy information for noun words. Additionally, we performed experiments using MSCOCO data in order to evaluate the performance of the proposed methodology.

▶ **Key words**: Deep Learning, Generative Adversarial Network, Text to Image Synthesis, Data Augmentation, WordNet

1. Multi Label Classification

Label Embedding Using AutoEncoder with Skip-Connections

- 연구 기간 : 2020.11. ~ 2021.04.
- 역할 : 주저자
- 사용 언어 및 프레임워크 : Python, Keras, Tensorflow

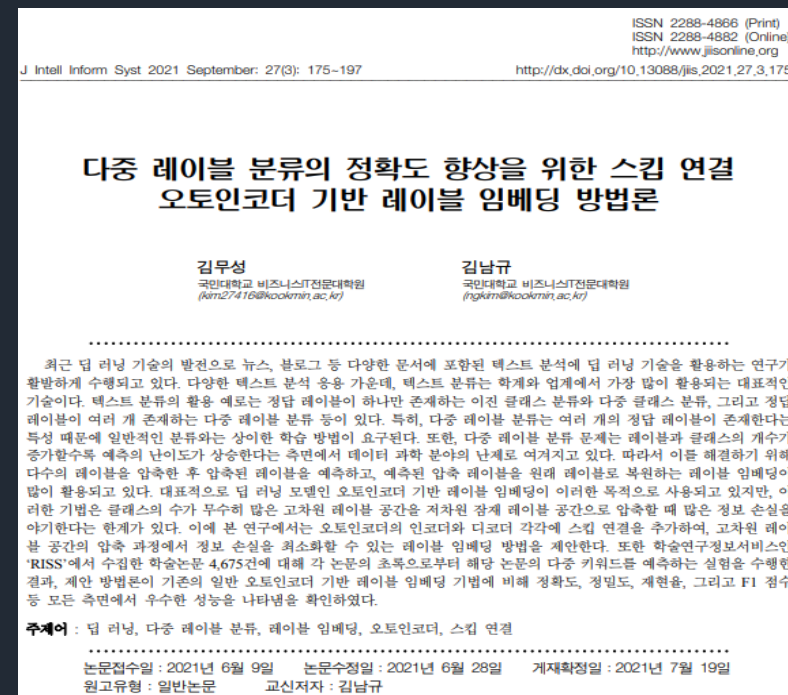
- 2021.06. 한국지능정보시스템학회 추계학술대회 발표 (최우수 논문상 수상)
- 2021.07. 다중 레이블 분류의 정확도 향상을 위한 레이블 임베딩 장치 및 방법 특허 출원
- 2021.07. 대학원생 논문기반 지식재산권 창출 지원 프로그램 공모전 (우수상 수상)
- 2021.09. Journal of Intelligence and Information Systems (KCI, To be published)

• 연구 문제

- ✓ 다중 레이블 분류는 레이블과 클래스의 개수가 증가할수록 예측 난이도가 증가함
- ✓ 이러한 문제를 해결하고자 했던 기존의 레이블 임베딩 역시 클래스의 개수가 증가할수록 정보 손실이 증가한다는 한계가 있음

• 연구 제안 내용

- ✓ 스킵 연결을 활용하여 정보의 손실을 최소화하는 방법을 제안함
- ✓ 스킵 연결 오토인코더 기반 레이블 임베딩 방법론을 제안함



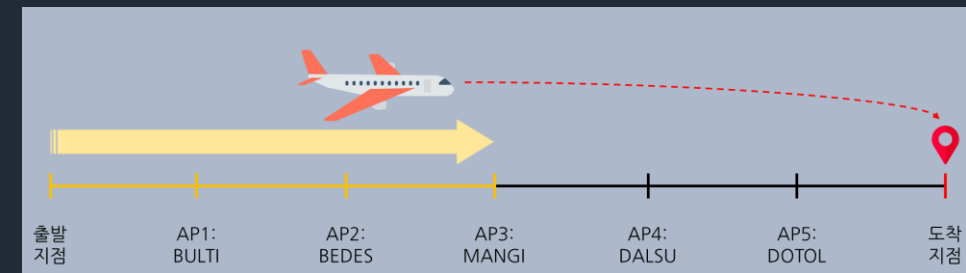
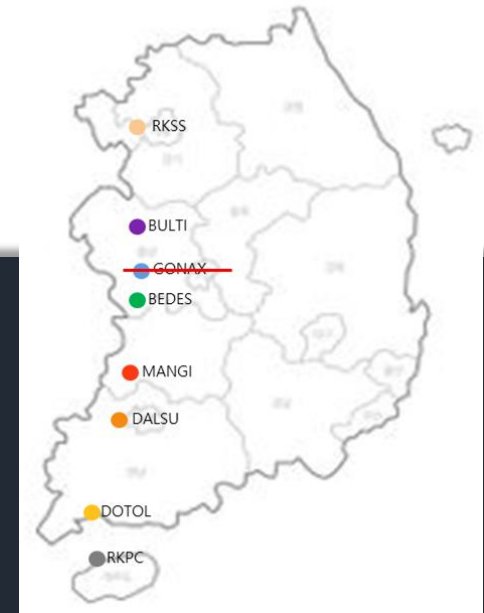
- 연구 기간 : 2021.05. ~ 2021.08.
- 역할 : 주저자
- 사용 언어 및 프레임워크 : Python, Pytorch
- 2021.08. ~ Paper Writing

- 연구 문제
 - ✓ 객체 인식 모델의 성능을 향상시키기 위해서는 많은 양의 데이터가 필요함
 - ✓ 기존의 이미지 데이터 증강은 객체의 정보를 훼손할 가능성이 높음
- 연구 제안 내용
 - ✓ 스타일 전이를 사용하여 객체 정보를 최대한 보존함
 - ✓ 스타일 전이 기반의 Two-Stage 이미지 증강 방법론을 개발함

2. Project 항공 빅데이터 및 다층 신경망 모델을 활용한 실시간 3D 운항분석 및 예측 플랫폼 개발

- 발주 기관명 : 스마트포트테크놀로지
- 연구 기간 : 2020.07. ~ 2021.02. (8개월)
- 프로젝트 개요
 - Part1. 항공 도메인 서버이
 - Part2. 항공 데이터 EDA 및 Feature Engineering
 - Part3. 예측 모델 개발(DNN, LSTM)

- 총 3가지 파트로 이루어진 시스템 중 Part2의 **파트장** 담당
- Part2. 항공 데이터 EDA 및 Feature Engineering
 - ✓ EDA를 통해 **운항 경로, 회항, 센서 오작동** 등의 정보 추출
 - ✓ 소요 시간 및 지연 시간 등의 **파생 변수 생성**
- 실시간 운항 지연 정보를 활용한 최종 운항 소요 시간 예측 모델 개발
- 실시간 운항 정보(지연 시간, 고도, 속도, 방향, 위도, 경도 등)만 사용
- 기존의 예측 모델과 딥러닝 기반 예측 모델 비교



2. Project 에스테틱 관리를 위한 진단 및 관리 후 사진 생성 모듈 구현

- 발주 기관명 : 약손명가 헬스케어
- 연구 기간 : 2021.05. ~ 2021.08. (4개월)
- 프로젝트 개요
 - Part1. 관리 전후 사진 데이터 구축
 - Part2. 데이터 전처리 및 고해상화
 - Part3. 관리 후 예측 이미지 생성(StyleGAN)

- 총 3가지 파트로 이루어진 시스템 중 Part2와 Part3의 **파트장** 담당
- Part2. 데이터 전처리 및 고해상화
 - ✓ 이미지를 정방형의 크기로 Crop 수행
 - ✓ 고화질 이미지 생성을 위해 고해상화 수행
- Part3. 관리 후 예측 이미지 생성
 - ✓ StyleGAN을 활용하여 관리 전과 관리 후의 벡터 추출
 - ✓ 벡터 연산을 통해 관리 후 이미지 생성



2. Project 지능형 사회안전 시스템 구축을 위한 멀티모달 딥러닝 기반 심층 상황 이해 모델 개발

- 발주 기관명 : 한국연구재단
- 연구 기간 : 2021.07. ~ 2022.02. (8개월)
- 프로젝트 개요
 - Part1. 텍스트-이미지 멀티모달 데이터셋 구축
 - Part2. 텍스트와 이미지로부터 의미 벡터 추출
 - Part3. 각각의 의미 벡터를 활용한 심층 상황 벡터 생성

- 총 3가지 파트로 이루어진 시스템 중 Part2와 Part3의 **파트장** 담당
- Part2. 텍스트와 이미지로부터 의미 벡터 추출
 - ✓ BERT를 활용하여 문장 벡터 추출
 - ✓ 이미지 캡셔닝과 BERT를 활용하여 이미지의 텍스트화 및 의미 벡터 추출
- Part3. 각각의 의미 벡터를 활용한 심층 상황 벡터 생성
 - ✓ 텍스트에만 나타난 벡터 추출
 - ✓ 이미지에만 나타난 벡터 추출
 - ✓ 텍스트와 이미지 모두 나타난 벡터 추출





감사합니다