**2023-1 인공지능과딥러닝 중간발표 Script**

**P1  
안녕하십니까, 인공지능과 딥러닝 6조 발표를 맡은 지민우입니다. 저희 조는 이번 Term project 주제로, 수기로 작성된 수식을 인식하는 OCR을 개발하고자 합니다. 발표를 시작하겠습니다.**

**P2  
광학 문자 인식 (Optical character recognition, OCR)이란, 사람이 손으로 작성하거나 인쇄된 문자 이미지를 획득하여 기계가 처리할 수 있는 Machine-encoded format으로 변환하는 기술입니다. 아래의 그림에서 보시듯, OCR 기술을 통하여 다양한 형태의 글자 데이터들을 컴퓨터가 처리할 수 있는 Text로 변환할 수 있습니다. 이러한 OCR 기술을 통해, 인쇄된 문서를 파일로 변환하는 과정을 자동화하거나, 텍스트 데이터를 검색 가능한 데이터로 변환할 수 있습니다. 이러한 장점을 지닌 OCR 기술은 정보자산의 디지털화에 있어서 중요한 기술입니다.**

**OCR 기술의 예시로, Papago의 이미지 번역 서비스를 들 수 있겠습니다. Text를 사진으로 촬영하면 Papago가 텍스트를 인식하여 번역하는 과정을 자동적으로 한 번에 수행합니다. 나아가, OCR 기술은 주민등록증, 사업자등록증 등 행정문서 처리 자동화에 적용되고 있으며, 물류 처리 시스템에도 활용되는 등 다양한 분야에 확대되어 적용되고 있습니다.**

**P3  
OCR은 피처 추출과 피처 분류로 주요 2 가지 과정을 통하여 이루어집니다. 오른쪽 그림에서 보시듯, 로드 된 이미지 데이터에 전처리를 수행하여 Labelling을 수행하고, Pattern matching을 통하여 수식을 인식하여 출력하는 방식으로 OCR이 수행됩니다. 특히, 저희 팀은 OCR 중 수학 기호를 인식하는 Mathematical symbol OCR에 관심을 가졌습니다. 일정한 글꼴을 갖는 Printed된 수식을 인식하는 상업용 OCR 프로그램은 비교적 우수한 성능을 확보하여 교육 현장 등 다양한 분야에 활용되고 있습니다. 한편, 수기로 작성한 수식을 인식하는 Handwritten mathematical symbols OCR의 경우, 인식의 정확도가 높지 않아 상업적 활용이 부진한 실정입니다. 수기로 작성한 수식의 경우, 개개인의 필체에 따른 차이, 이미지의 회전을 비롯한 여러 Noise가 존재하기 때문에 이미지 전처리 과정을 보다 정교하게 수행하여야 합니다.**

**P4.  
이어서 관련 연구에 대하여 말씀드리겠습니다. 아래의 연구에서는 수기로 작성한 수식을 인식하는 OCR을 개발하여, 이를 아동용 교육용 로봇에 적용하였습니다. 통상적인 OCR 과정을 토대로, 이미지 데이터에 전처리를 수행하고, CNN을 활용하여 기호를 인식하였으며, 출력 결과를 Equation solver에 대입함으로써 사람이 작성한 수식의 결과를 출력하는 프로그램을 개발하였습니다. 해당 연구에서는 0-9까지의 숫자 및 단순한 사칙연산과 같은 Basic expression에 집중하여 인식을 수행하였으며, 보다 복잡한 수식을 인식하는 과정을 포함하고 있지 않습니다. 특히, 해당 연구는 MATLAB을 사용하여 OCR을 구현하였습니다. 저희 팀은 해당 연구 내용을 Python을 사용하여 보완하고자 합니다.**

**P5  
먼저 예측의 정확도를 높이기 위해 이미지 전처리가 필요합니다. 수식영역을 찾기 위해 수식을 감싸는 BOX를 이용합니다. Opencv를 사용해 박스의 모서리를 검출하고 이미지를 크롭합니다. 그 후 grayscale로 변환하여 전처리를 마무리합니다.**

**P6  
이렇게 전처리가 된 이미지를 사용해 character를 하나씩 찾아냅니다. 그런데 = 같이 한 character 이지만 떨어져 있는 경우 서로 다른 character 라고 인식하는 경우가 있습니다. 이를 해결하기 위해 opencv filter의 erosion 기능을 사용해서 글자의 폭을 늘려 하나로 인식할 수 있도록 처리합니다. 이렇게 글자를 찾아내면 숫자 1처럼 가로세로 비율이 정사이즈가 아닌 경우를 처리하기 위해 흰색배경을 패딩으로 넣어 정사각형 사이즈로 맞춰줍니다.**

**P7  
학습은 Kaggle의 Handwritten math symbol dataset을 사용했습니다. 총 데이터 개수가 약 37만개 정도 있었고 9:1 비율로 트레이닝과 테스트 셋을 나눴습니다. CNN model을 사용해서 학습을 진행했고 학습이 끝난 결과 정확도가 약 98% 가 나왔습니다.**

**P8  
먼저 로컬 환경에서 트레이닝과 테스트를 진행했습니다. 오른쪽 보이는 환경에서 학습을 진행했을 때 학습에 걸리는 시간은 약 45분 정도가 걸렸습니다. 손으로 글씨를 써서 테스트를 해보면 다음처럼 결과가 나오게 됩니다.**

**P9**

**Quantitative evaluation을 진행해보았을 때는 40개의 이미지를 테스트 해봤을 때 평균 정확도는 약 75% 정도가 나왔습니다.**

**P10  
이번 프로젝트에서는 image preprocessing을 통해 노이즈를 제거함으로써 인식 정확도를 향상시킬 수 있었습니다. 수식 인식 과정에서 일부 기호는 서로 다른 객체로 인식되는 issue가 있었고 opencv의 필터를 통해 해결하였습니다. 그러나 분수 등과 같이 2차원적으로 작성된 수식, 혹은 root처럼 식을 감싸고 있는 수식은 적용할 수 없다는 한계가 있습니다.**

**향후 개선점으로는 2차원적으로 작성된 수식은 이미지 내의 좌표를 사용해 인식하고 처리하는 과정이 필요하다고 생각되며, 필체 차이에 따른 정확도 감소는 모델의 개선과 학습 데이터의 추가, 이미지 전처리 과정 보완이 필요하다고 생각됩니다. 그렇게 되면 향후 지수나 미적분 등이 포함된 다양한 수식을 인식할 수 있을 것으로 생각합니다.**

**P7  
저희가 준비한 자료는 여기까지입니다. 경청해주셔서 감사합니다.**