OCR 모델링 정의서

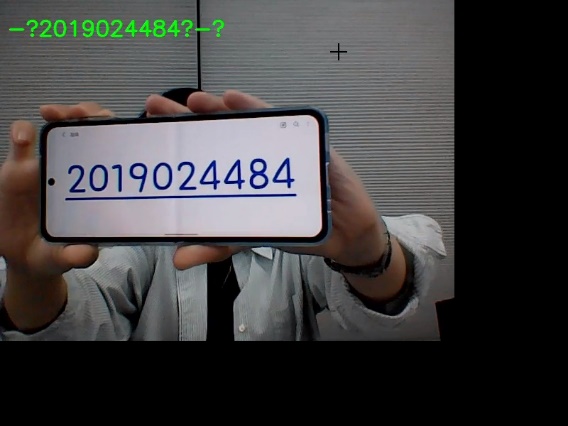
1. **OCR 모델 설명 (간략히)**

* Tesseract OCR : GOOGLE에서 개발된 무료 오픈소스 OCR(광학 문자 인식) 엔진  
  LSTM 기반의 딥러닝 이용 문자 인식 수행함  
  다양한 언어를 지원하고, 오픈 소스로 공개되어 가장 쉽게 자료를 찾을 수 있어 선택
* easyOCR : JaidedAI에 의해 개발된 오픈 소스 OCR 도구  
  pytorch 딥러닝 프레임워크를 기반으로 하여 정확도가 향상된 모델이고(가장 중요) 일반 텍스트 뿐만 아니라 날짜, 시간, 숫자, 화폐, 주소 등 다양한 형식의 텍스트 인식이 가능했으며 개발자 커뮤니티가 활발한 점을 장점으로 꼽아 선택
* paddleOCR : PaddlePaddle에서 제공하는 오픈소스 OCR 도구  
  이미지 전처리, 텍스트 검출, 텍스트 인식 등 모듈들 제공해 따로 이미지 전처리가 필요없고 정확도가 높으며 경량화된 모델로 실시간 처리가 가능.

1. **모델 검증 (TEST, 모델 별, 글씨체 별)**
2. Tesseract OCR로 사진 OCR 실행  
   - basic tesseract 사용 : 성능이 별로 좋지 않음  
   - 사진에 대하여 전처리&텍스트 윤곽선 찾기 실행  
   텍스트 윤곽선을 찾기 위한 grayscale 변환, blurred 처리, canny로 모서리 탐지를 실 행한 이미지에서 findContours() 메소드를 이용해 윤곽선을 찾으려 했음. 하지만 텍스트만 있는 이미지가 아닌 배경이 포함된 이미지여서 윤곽선을 잘 찾지 못했음.

  
- 윤곽선 찾기 빼고 이미지 전처리만 실행  
grayscale, 노이즈 제거, 밝기 및 대비 조절, 이진화, 모폴로지 연산 처리를 실행. 어떤 처리를 했을 때 가장 잘 찾아냈는지 test했음. 이진화와 모폴로지 연산은 텍스트 자체를 날리게 되는 경우가 있어 (조명, 자리 등에 따라 설정값을 바꿔야 했기에 모든 사진에 같은 처리를 입힐 수 없었음) 제외하고 최종적으로 grayscale 변환, 노이즈 제거, 밝기 및 대비 조절만 선택

1. Tesseract OCR 웹캠 연결 실시간 OCR 실행  
   이미지 전처리 결과로 텍스트를 얼추 찾아냈지만, 텍스트가 멀리 떨어진 경우(약 20츠 이상) 성능이 많이 떨어지는 것을 확인하고 더 성능이 좋은 모델을 찾아봄

 > 이 정도 거리에서만 탐지

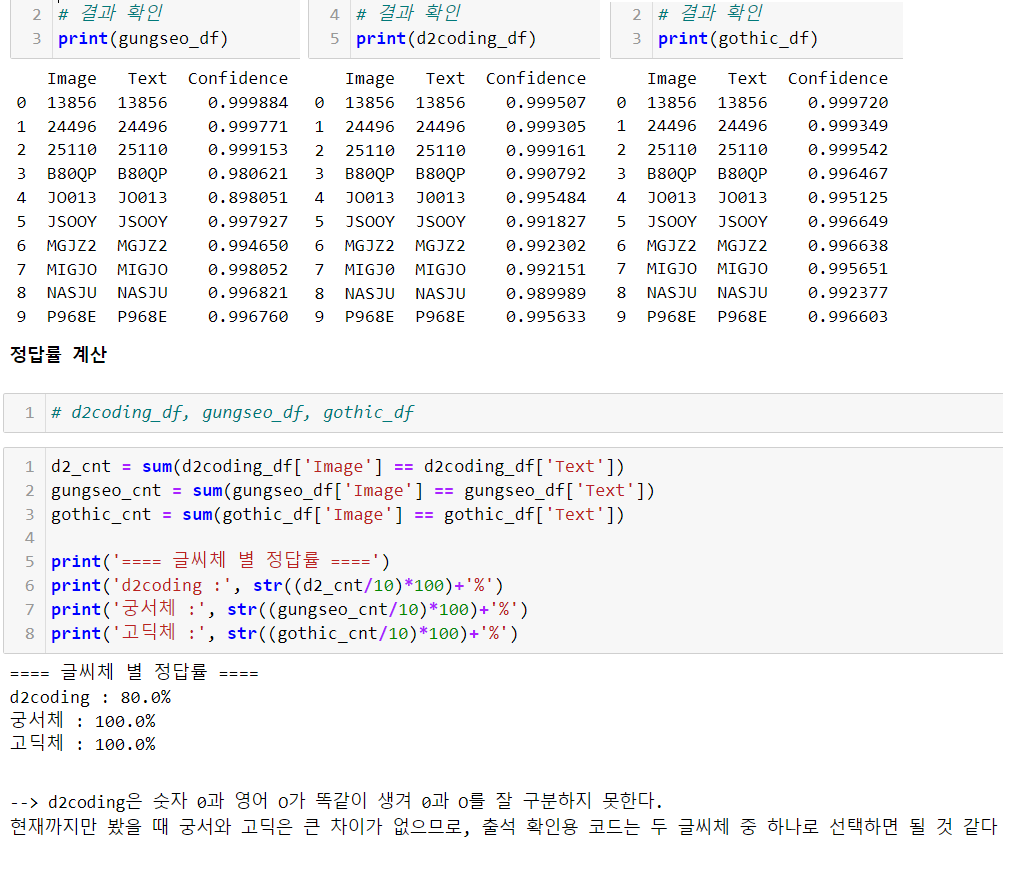
1. 정확도 향상 모델 easyOCR 사진 ocr 실행  
   위에서 얻은 전처리 단계인 grayscale 변환, 노이즈 제거, 밝기 및 대비 조절 실행 후 ocr을 실행한 결과 tesseract에서는 찾지 못하던 텍스트도 찾아내는 것을 발견
2. easyOCR 웹캠 연결 실시간 ocr 실행  
   실시간으로 웹캠을 연결했을 때 어느 정도 거리가 멀어져도 텍스트를 잘 찾아내지만, 모델 크기가 커서 실시간 화면에 로딩이 걸리고 버벅거리는 현상을 확인
3. 경량화 모델 paddleOCR  
   - 경량화된 모델이라고 알려진 paddleOCR 모델로 웹캠을 연결한 실시간 OCR 실행  
   - 모델 내부에서 이미지 크기 조정, 노이즈 제거, 이진화 작업 등을 자동으로 수행하므로 이미지 전처리 과정 삭제  
   - 텍스트와 카메라 거리가 1미터 이상으로 멀어져도 텍스트 인식을 잘 했고, 이전 모델들(tesseractOCR, easyOCR)보다 정확도도 높았음.   
   - 실시간으로 모델을 돌려도 크게 버벅거림이 없어 paddleOCR 모델을 최종으로 선택함  
   의류, 사람, 반사안전피복, 스크린샷이(가) 표시된 사진

   자동 생성된 설명

> 처리 속도 0초대, 거리 멀어도 텍스트 탐지 가능, 실시간 버벅거림 없음

1. **성능지표 및 결과 (최종 사용 결과)**

* tesseractOCR : 정확도 낮음, 거리 멀어지면 텍스트 탐지 불가능, 실시간 연결 가능
* easyOCR : 정확도 높음, 거리 멀어지면 텍스트 탐지 불가능, 실시간 연결 불가능
* paddleOCR : 정확도 높음, 거리 멀어져도 텍스트 탐지 가능, 실시간 연결 가능
* 텍스트 글씨체 선택 > 궁서체 선택



1. **결과 해석 및 분석 (왜 선택?)**

제일 처음 선택했던 tesseractOCR 모델은 이미지 전처리를 했음에도 불구하고 정확도가 낮고, 속도도 느렸다. 그후 선택한 easyOCR 모델은 속도는 빠르고 정확도도 높았지만 텍스트 거리가 멀어지면 잘 탐지를 하지 못했고, 실시간 처리를 돌렸을 때 모델이 무거워 버벅거리는 현상이 발생했다. 마지막으로 선택한 경량화된 모델 paddleOCR은 이미지 전처리를 따로 하지 않았음에도 1.5m 거리까지는 텍스트 인식이 가능했고, 실시간 처리에서도 버벅거림 없이 높은 정확도를 보여주었기 때문에 paddleOCR 모델을 선택했다.