

딥러닝을 활용한 알약 분석 어플리케이션

김민형, 김홍준, 허민혁
인천대학교 정보통신공학과

kmh951210@naver.com, 96hjun@naver.com, alsgur5731@naver.com

Pill analysis application utilizing deep learning

Min-hyeong Kim, Hong-Jun Kim, Min-hyeok Heo
Dept. of Information and Communication Engineering,
Incheon National University

요 약

예전보다 약과의 접촉이 많아지고 쉬워지면서 집안에 방치되거나 복용을 잊어버린 약들이 더 많아졌다. 이런 약들을 다시 먹으려 할 때는 특별한 주의가 필요하다. 특히나 임산부, 어린이등 취약 계층은 더욱 주의가 필요하다. 이들 모두가 간단하게 이용할 수 있게 사진을 찍으면 약의 이름과 효능, 부작용, 복용방법 등에 대한 정보를 알 수 있는 어플리케이션을 개발했다.

1. 서론

24시 편의점에서도 약을 구매할 수 있을 정도로 현대인들의 알약과의 접촉은 보다 많아지고 손쉬워졌다. 그만큼 약의 종류도 많아지고 가져오는 효과나 부작용도 늘어나고 있어 어느 때 보다 약을 복용할 때 주의가 필요하다.

현재 알약 정보를 확인할 수 있는 사이트는 약학정보원, 드러그인포, 네이버 알약사전 등이 있다. 이런 사이트들을 보면 알 수 있듯이 약 정보를 검색하는데 있어 약에 적힌 문자, 약의 색 정보, 모양, 제형, 분할선등을 직접 선택하는 방법으로 약의 정보를 알 수 있다. 이는 약의 접촉이 상대적으로 많고 더 위험한 취약계층 이 사용하기에는 불편하고 어려운 점이 많다.



사진1(사진 검색 방법 좌: 약학정보원, 우 : 드러그인포)

따라서 모두가 손쉽게 사용할 수 있게 사진으로

알약 정보를 확인할 수 있는 어플리케이션을 개발했다.

2. 관련연구

2.1 웹 크롤링

스크레이핑(scraping)이라고도 불리며 웹 페이지를 가져온 후 데이터를 추출하는 기술이다. 인터넷 상에는 엄청난 양의 정보가 있으며 이를 직접 하나하나 추출하는 과정은 불가능에 가깝다. 따라서 정보를 필요로 할 때 많이 사용된다. 웹은 기본적으로 html형태로 되어있으며 정형화된 형태로 관리되어지므로 규칙이 생기게 된다. 크롤링은 이러한 규칙을 분석해 정보를 추출하며 대표적으로 python, ruby 등이 있다. 본 논문에서는 python의 beautifulsoap를 이용하였다.

2.2 데이터 전처리

분석을 위하여 데이터를 그에 맞게 변경하는 과정을 데이터 전처리(Data Processing)라고 한다. 좋은 분석 결과를 얻기 위해선 정확한 데이터가 입력되어야 하므로 데이터 전처리 과정은 매우 중요하고 필수적인 과정 중 하나이다. 데이터 전처리에는 다양한 방법이 있으며 숫자로 변형, 정규화(normalize), 필터링(filtering) 등이 있다. 본 논문에서는 이미지의 명암, 회전, 반전, 2진처리 등의 과정으로 학습할 파일을 준비하고 이미지 파일을 다차원 배열로 변환하는 전처리 과정을 진행하였다.

2.3 keras/cnn

케라스(Keras)는 Python으로 작성된 오픈 소스 신경망 라이브러리이다. MXNet, Deeplearning4j, 텐서플로, Microsoft Cognitive Toolkit 또는 Theano 위에서 수행할 수 있다. 딥 신경망과의 빠른 실험을 가능케 하도록 설계되었으며 최소한의 모듈 방식의 확장 가능성에 초점을 둔다. 더 높은 수준의 더 직관적인 추상화 집합을 표현함으로써 백엔드 과학 컴퓨팅 라이브러리임에도 불구하고 신경망을 구성하기 쉽게 만들어준다.

합성곱 신경망(Convolutional neural network, CNN)은 시각적 이미지를 분석하는 데 사용되는 깊고 피드-포워드적인 인공신경망의 한 종류이다. 딥 러닝에서 심층 신경망으로 분류되며, 시각적 이미지 분석에 가장 일반적으로 적용된다. 또한 공유 가중치 구조와 변환 불변성 특성에 기초하여 변이 불변 또는 공간 불변 인공 신경망(SIANN)으로도 알려져 있다. 이미지 및 비디오 인식, 추천 시스템, 이미지 분류, 의료 이미지 분석 및 자연어 처리에 응용된다.

CNN은 네트워크가 기존 알고리즘에서 수작업으로 제작된 필터를 학습한다. 피쳐 디자인에 대한 사전 지식과 인간 노력과의 독립성은 CNN의 주요한 장점이다.

본 논문에서는 모델 훈련과 평가를 동시에 진행하기 위해 keras와 cnn을 사용하여 알약의 이미지를 학습전용과 테스트 전용으로 구분하여 학습시키는 방법을 사용하였다.

모델은 계층을 선형으로 쌓는 sequential 모델을 사용하였고, 내부 활성화 함수로는 relu 라는 활성화 함수를 사용 하였다. relu 활성화 함수는 0 보다 큰값일 경우 1을 반환하는 sigmoid와는 다르게 0보다 작은 값이나올경우 0을 반환하고 0보다 큰값이 나온 경우 그 값을 그대로 반환한다. 따라서 정확도를 올리기 위해 내부 활성화 함수에 relu함수를 적용하고 output layer 에서만 sigmoid 함수를 적용했다.

모델의 입력 형태는 필터와 커널, 경계처리 방법을 정의해 더해주는 Convolution2D, 사소한 변화가 영향을 미치지 않도록 수직, 수평 축소비율을 설정해 주요값만 뽑아 크기가 작은 출력 영상을 만들어주는 MaxPooling2D, 과적합 방지를 위해 학습시에 지정된 비율만큼 임의의 입력 뉴런

을 제외시키는 Dropout 그리고 함수를 통해 위에서 다뤄진 2차원 자료들을 Dense(전결합층)에 전달하기 위한 1차원 자료로 변경 시켜주는 Flatten을 활용하였다.

2.4 안드로이드 스튜디오

안드로이드 스튜디오(영어: Android Studio)는 안드로이드 및 안드로이드 전용 어플(앱) 제작을 위한 공식 통합 개발 환경 (IDE)이다. 과거 이클립스 기반의 안드로이드 개발 툴인 Android Development Tool의 주요 빌드 시스템은 아파치 앤트였으나 공식 안드로이드 스튜디오는 gradle빌드 시스템을 사용하고 있다. SDK 매니저(SDK manager)를 통해서 안드로이드 버전별 리소스나 C 및 C++ 을 빌드할 수 있는 네이티브 어플리케이션 제작 등 다양한 개발도구들을 추가적으로 지원한다. 프로그래밍 언어로는 자바 이외에도 코틀린을 지원한다. 이처럼 모델을 구현할 어플리케이션을 디자인하기 위해 본 논문에서는 Android Studio(java)를 사용하였다.

준비 단계에서 미리 알약에 대한 이미지와 정보를 크롤링을 통하여 가져왔다. 그 데이터들을 어플리케이션 내부에 저장을 해두었다.

모델 초기화 (init model) 부분에서 Assets 폴더에 만들어준 라벨 문서(label.txt)에서 알약 이름 리스트들을 읽어온다. 그 후 Firebase에 등록된 모델자동업데이트, 모델 소스 등록, 모델 검출기 생성 등의 작업을 한다.

어플리케이션이 실행되면 새로운 사진을 찍어서 입력할 것인지 기존의 기기 내부에 있던 사진을 쓸 것인지 정할 수 있게 했다. 입력 이미지를 정한 후에 예측을 시작하게 되면 내부적으로 모델의 입력형식과 출력형식을 지정하고(본 논문에서는 입력 형식(1,50,50,3), 출력 형식 (1,15)으로 지정) 모델 입력데이터 (본논문에서는 50x50 이미지 사이즈 지정)를 이미지를 이용해 생성한다. 학습시킨 모델 파일 (tflite)를 이용해 입력 이미지를 분석하여 가장 비슷한 알약을 출력한뒤 크롤링하여 저장해둔 알약 이미지와 정보를 매칭하여 출력한다.

3. 설계 및 구현

3.1 서비스 구성도 및 흐름

사용자가 스마트폰 카메라로 이미지를 획득하면 스마트폰에 탑재된 tflite 모델이 이미지 데이터를

비교한다. 가장 일치하는 이미지 데이터를 찾으면 해당 데이터 정보를 전달해 준비된 약 사진과 정보를 스마트폰 화면에 출력해 준다.(사진2 구성 흐름 참조)

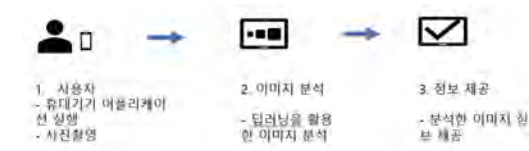


사진2(구성 흐름)

3.2 기능 처리도 및 흐름도

APP 실행하여 카메라/갤러리 옵션을 선택해 옵션에 따라 확보한 이미지를 딥러닝 모델에 입력해 이미지를 분석한다. 분석된 결과를 약정보와 매칭해 최종적으로 결과를 화면에 출력시켜준다. (사진3 기능처리 흐름 참조)

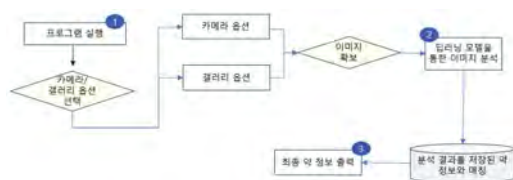


사진3(기능처리 흐름)

3.3 모델 테스트

모델의 평가를 용이하게 하기 위해 해당 모델을 학습할 때 학습전용과 테스트 전용 파일을 분리해서 실행했다. 알약 이미지의 특성상 다른 이미지에 비해 단순하고 이미지 데이터가 많이 없어 구분을 지어도 원본파일은 같은 출처의 이미지를



사용할 수밖에 없어 모델 학습을 완료한 후 해당 모델을 테스트함으로써 문제점을 파악했다.

사진4(모델 테스트 좌 : 테스트 성공 우 : 테스트 실패)
해당 모델에서는 학습된 이미지 파일로 직접 평가했을 때 100%의 정확도가 나왔지만 직접 찍은 사진으로 시도했을 때는 대부분의 테스트가 실패 배경 제거와 모델 보완 등의 문제점을 확인할 수 있었다.

3.4 구현

데이터 수집 과정은 Python의 beautifulsoup 라이브러리를 이용하여 진행하였으며 네이버 알약 사전에 있는 이미지들을 자동으로 추출(크롤링)하여 사진을 수집하였다. 수집한 이미지들은 Python cv2 라이브러리를 이용하여 OpenCV를 통해 데이터 전처리를 하였다. 그 후에 처리가 완료된 데이터들을 저장하고 딥러닝 과정에서 학습에 적절하지 않다고 생각되는 이미지들은 삭제하였다.

딥러닝 모델 생성 과정에서는 내부 활성화 함수로 0보다 작은 값이 나오면 0을 반환하고 0보다 큰 값이 나오면 그 값 그대로 반환하는 relu 함수를 이용하였다. 출력층에서만 sigmoid 함수를 적용하였으며 처음 softmax 함수를 적용하였을 때 80%대였던 정확도를 90% 이상의 정확도로 향상시킬 수 있었다.

학습시킨 모델을 스마트폰 어플리케이션에 올리기 위해선 스마트폰 환경에 맞춰 경량화 시킬 필요가 있었다. 이 요구에 맞춰 모델 파일을 h5파일을 tflite 파일로 변환시켜 경량화 시키고, 파이썬 베이스의 머신러닝 모델에 업로드 하여 어플리케이션에서 사용할 수 있게 하였다. 변환하는 과정에서는 Google Colab을 이용하였다.

안드로이드 스튜디오에서는 어플리케이션 초기에 갤러리에서 선택하거나 사진 촬영으로 이미지를 가져올 수 있도록 했고 해당 입력 이미지를 탑재된 tflite 모델에서 분석할 수 있도록 하여 가장 정확도가 높은 알약에 대한 정보를 출력해줄 수 있도록 하였다.

모델 초기화 (init model) 부분에서 Assets 폴더에 만들어준 라벨 문서(label.txt)에서 알약 이름 리스트들을 읽어오게 하였다. 이미지를 입력하여 모델 분석을 완료하면 정확도가 높은 상위 3개의 알약이 출력되며 상위 3개 알약의 이름과 정확도를 가장 위에 보여주며, 이미지와 알약에 대한 정보는 가장 정확도가 높은 알약에 대해서만 출력한다.

해당 프로젝트에서는 알약에 관한 이미지와 정보를 미리 크롤링해서 저장해 두었고, 추후 테스트가 완료된 후엔 데이터베이스 서버와 연결해 해당 정보를 가져오는 과정이 추가되어야 원활한 정보 출력이 가능할 것이다.



사진5(좌 : 어플리케이션 초기화면 중, 우 : 탁센,프로페시아 알약 결과 출력화면)

4. 결론

본 논문에서는 일상 생활에서 촬영된 알약 이미지를 인식하여 정보를 제공하는 어플리케이션을 제안한다. CNN을 이용하여 알약 이미지의 MEAN Filter를 이용한 이진화, 감마값 조절, 회전 등을 조절하여 알약의 다양한 모습들을 학습하게 하였다. 완성된 모델을 tflite파일로 변환하여 모바일에서도 사용할 수 있게 하였으며 배경 제거와 이진화 부분에서 좀 더 좋은 결과를 얻고 모델의 구조를 개선한다면 인식률은 계속 향상될 것이며 실생활에서 유용하게 사용될 것이다.

참고문헌

김대욱, 스마트폰으로 촬영된 알약 영상의 글자 및 형상 인식 방법, 공학석사 학위 논문, 2017.8

고경철, 박세혁, 고국원, 알약 외관 검사용 영상처리 시스템, 한국정밀공학회 학술발표대회 논문집, 29-30, 2008년

구건서, 알약 인식을 위해 색 특징정보를 이용한 CBIRS/TB, 한국컴퓨터정보학회논문지, 19(2), 49-56, 2014년

[본 논문은 과학기술정보통신부 정보통신창의인재 양성사업의 지원을 통해 수행한 ICT멘토링 프로젝트 결과물입니다.]