Journal of Biomedical Engineering Research 43: 441-447 (2022)
http://dx.doi.org/10.9718/JBER.2022.43.6.441

학 술 논 문

음각 정보를 이용한 딥러닝 기반의 알약 식별 알고리즘 연구

이선민[1] · 김영재[2] · 김광기[3]*

[1]가천대학교 간호대학 간호학과, [2]가천대학교 의과대학 의공학교실
[3]가천융합의과학원 융합의과학과

Pill Identification Algorithm Based on Deep Learning Using Imprinted Text Feature

Seon Min Lee[1], Young Jae Kim[2] and Kwang Gi Kim[3]*

[1]Department of Nursing, College of Nursing, Gachon University
[2]Department of Biomedical Engineering, Gill Medical Center, College of Medicine, Gachon University
[3]Department of Health Sciences & Technology, Gachon Advanced Institute for Health Sciences & Tecnology (GAIHST), Gachon University

(Manuscript received 30 November 2022 ; revised 22 December 2022 ; accepted 23 December 2022)

Abstract: In this paper, we propose a pill identification model using engraved text feature and image feature such as shape and color, and compare it with an identification model that does not use engraved text feature to verify the possibility of improving identification performance by improving recognition rate of the engraved text. The data consisted of 100 classes and used 10 images per class. The engraved text feature was acquired through Keras OCR based on deep learning and 1D CNN, and the image feature was acquired through 2D CNN. According to the identification results, the accuracy of the text recognition model was 90%. The accuracy of the comparative model and the proposed model was 91.9% and 97.6%. The accuracy, precision, recall, and F1-score of the proposed model were better than those of the comparative model in terms of statistical significance. As a result, we confirmed that the expansion of the range of feature improved the performance of the identification model.

Key words: Pill identification, Deep learning, Imprinted text, Keras OCR, CNN

I. 서 론

입원환자의 처방은 용량 및 용법, 환자의 상태, 보험 적용 등이 변경되어 취소될 수 있고, 이때 처방되었으나 투약되지 않은 약은 병동 약국으로 회수된다[1]. 회수된 알약은 약사의 육안으로 재분류하여 반납 처리를 완료하고, 이때 색과 모양이 유사한 알약 분류는 혼동을 유발하여 반납된 약을 재사용하는 과정에서 투약 과오의 위험이 발생한다[2]. 또한 약국에서는 의약품 공급, 복약지도, 약력관리 등의 기타 업무 시간 및 인력 자원의 낭비가 발생한다[3]. 이러한 문제는 알약 정보(feature)의 획득 및 식별을 자동화하여 해결 가능하며, 그 수단으로 인공지능을 활용한 연구들이 다수 진행되어 왔다.

인공지능의 단순 식별은 색상, 모양과 같은 알약의 고유 정보에 기반하여 수행되었다. 그 예시로, 2017년 Wang 등은 7,000장의 이미지에서 알약의 정보를 획득한 후 1,000종의 알약을 식별하였고, 그 결과 mAP scores는 0.328이었다. 제안된 방법은 딥러닝을 사용했음에도 불구하고 촬영 환경의 편차가 심하여 식별 성능이 다소 낮았다[4]. 2019년 Delgado 등은 여러 Convolutional Neural Network(CNN) 모델을

*Corresponding Author : Kwang Gi Kim
38-13, Docjeom-ro 3beon-gil, Namdong-gu, Incheon, 21565, Republic of Korea
Tel: +82-32-458-2770
Fax: +82-32-460-2361
E-mail: kimkg@gachon.ac.kr
본 연구는 중소벤처기업부의 맞춤형 기술파트너 지원사업(RS-2022-00165316)과 경기도의 경기도 지역협력연구센터 사업[GRRC-가천2020(B01), AI기반 의료영상분석]과, 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음(IITP-2022-2017-0-01630).

www.kci.go.kr

사용하여 1,000종의 알약을 식별하고 모델 간 성능을 비교한 결과, 가장 우수한 모델의 정확도가 77.10%로 확인되었다[5]. 2020년 Yang-Yen Ou 등은 2만여장의 이미지에서 알약 612종을 검출하고 Inception-ResNet-V2 모델을 이용하여 식별하였으며, 식별 성능은 오류율(Error rate) 17.9%로 확인되었다[6]. 두 연구 모두 딥러닝 모델의 알약 식별 적합성을 확인하였다는 데에 의의가 있으나, 제안한 방법을 약무현장에 적용하기 위해서는 식별 성능이 더욱 향상되어야 한다.

앞의 방법들은 딥러닝 모델을 이용하여 알약을 식별하였다. 그럼에도 불구하고 모양, 색상 등의 정보가 유사하고 알약 종류가 많거나 촬영 환경의 편차가 클수록 식별에 한계가 있음을 확인하였다. 이러한 한계는 알약 식별에 활용되는 정보의 범위를 확대하여 해결할 수 있다.

글자 정보는 알약의 표면에 각인되거나 프린트된 문자로, 모양 및 색상과 함께 알약 식별에 활용할 수 있는 중요한 척도이다. 글자 정보를 이용한 알약 식별의 자동화는 앞서 Suntronsuk 등에 의해 발표된 바 있다. 2017년 Suntronsuk 등은 영상처리로 음각 글자와 프린팅된 글자를 분할하고, Tesseract OCR과 K-평균 군집화를 이용하여 알약을 식별하는 방법을 제안하였다. 음각과 프린트된 글자의 구분 없이 산출한 정확도는 57%, 음각 글자의 경우 글자 인식과 알약 식별의 정밀도는 각각 28%, 35%, 프린트된 글자의 경우 글자 인식과 알약 식별의 정밀도는 각각 66%, 75%로, 음각 글자에서의 성능이 현저히 낮음을 확인하였다[7]. 한편 같은 해에 김대욱은 음각 글자 정보에 모양 정보를 취합한 자동 식별 연구를 진행하였다. 글자 영역을 박스 형태로 바운딩(bounding)하여 획득하고, 직접 개발한 CNN 기반의 OCR을 학습하여 글자를 인식하였다. 이는 14가지 클래스로 임의 구별한 모양 정보와 함께 알약 식별에 사용되었고, 그 결과 글자 영역의 검출률 87.5%, 글자의 인식 정확도 55.4%, 글자와 모양 정보에 기반한 식별 정확도 58.4%를 확인하였다[8].

두 연구 모두 글자 정보를 활용하여 알약을 식별하였으나, 김대욱은 모양 정보까지 취합하여 Suntronsuk 등에 비해 향상된 식별 성능을 나타내었다. 또한 Tesseract OCR에 비해 노이즈가 있는 이미지에서의 글자 인식률이 더 우수한 CNN[9] 기반 OCR 모델을 직접 개발하였다는 점에서 의의가 있다. 그럼에도 불구하고, 그 성능의 수준은 제안된 방법을 실용화하기에 적합하지 않다. 또한 제안 방법은 영상처리를 이용하여 글자 영역을 획득하고 Zernike 모멘트를 계산하여 모양 정보를 획득하므로, 과적합의 위험이 높다[10]. 게다가 모양 정보와 글자 정보를 조합하고, 그와 일치하는 알약 종류를 도출하는 방법을 딥러닝으로 대체한다면 최종 식별 성능을 개선할 수 있다. 이러한 한계점을 극복하여 딥러닝 기술로 글자 정보와 모양, 색상 등의 이미지 정보를 자동으로 획득하고, 음각 글자의 인식률과 알약 식별률을 대폭 개선하는 연구가 필요하다.

이에 따라 본 논문은 딥러닝 기술 기반의 OCR로 글자 영역의 획득 및 정보 인식을 수행하고, 이미지 정보와 함께 알약을 식별하는 딥러닝 모델을 제안하고자 한다. 또한 글자 정보를 포함한 식별 모델과 글자 정보를 제외한 식별 모델을 비교 검증하고자 한다.

II. 연구 방법

1. 개발환경

실험 환경은 NVIDIA GeForce RTX 1050 TI(NVIDIA, Santa Clara, Calif) 그래픽 처리 장치와 2.90 GHz Intel® Core™ i7-10700(Intel, Santa Clara, Calif) CPU 및 32GB RAM으로 구성되었다. 모델 학습은 Python 3.7.0(Python Software Foundation, Wilmington, DE, USA) 및 Keras 2.1.3(Tensorflow) 프레임워크, 모델 평가는 Scikit-learn 1.0.2 프레임워크를 사용하였고, 이미지 전처리는 OpenCV 3.4.1 라이브러리를 사용하였다.

2. 데이터 수집

본 연구는 시중에 유통되고 있는 알약 중 정제(tablet) 형태의 알약 100건을 선정하였고, 그 중 알약의 모양에 따라 원형 정제 80건, 타원 및 장방형 정제 20건으로 나누어 구성하였다(Fig. 1). 알약 종류별 10장의 앞면 이미지를 수집하였으며, 식별 모델에 사용된 총 1,000장의 이미지에서 900장은 학습 데이터, 100장은 테스트 데이터로 구성하였고 테스트 데이터는 랜덤하게 선정하였다. 또한 음각 인식 모델은 식별 모델에서 나눈 테스트 데이터 100장을 사용하였다. 모든 데이터는 스마트폰으로 촬영된 24비트 컬러의 원본 이미지에서 알약 영역을 정사각형 형태로 수동 크롭하였다.

3. 음각 인식 모델

음각 글자 인식에 사용된 모든 이미지는 512*512픽셀, 평균 밝기는 0.75 이하로 하향 조정하였고(Fig. 2의 (b)), 음각 영역 외의 불필요한 배경과 그림자를 제거하기 위해 원형 정제와 타원 및 장방형 정제로 나누어 영상처리 기법을 적용하였다. 원형 정제의 경우, 허프변환[11]을 이용하여 알약 윤곽선에 가까운 원을 검출하고 검출 영역에 대한 마스크 이미지를 획득하였다(Fig. 2의 (c)). 타원 및 장방형 정제의 경우, canny edge 알고리즘[12]을 이용하여 알약 윤곽선을 검출하고, 모폴로지 연산[13]과 medianBlur 필터[14] 등을 적용하여 윤곽을 원본에 가깝게 보정하였다. 원형과 타원 및 장방형 모두 최종적으로 획득한 마스크 이미지와 컬러 이미지를 AND 연산으로 합성하였고(Fig. 2의 (d)), 합성된 이

www.kci.go.kr

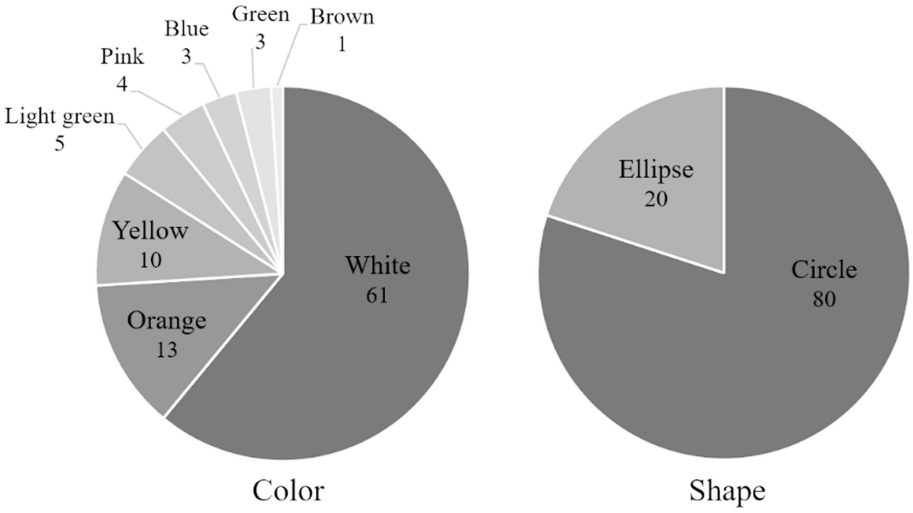

그림 1. 색깔과 모양에 기반한 데이터 구성의 원형 그래프
Fig. 1. Circle figure of data proportions by color and shape

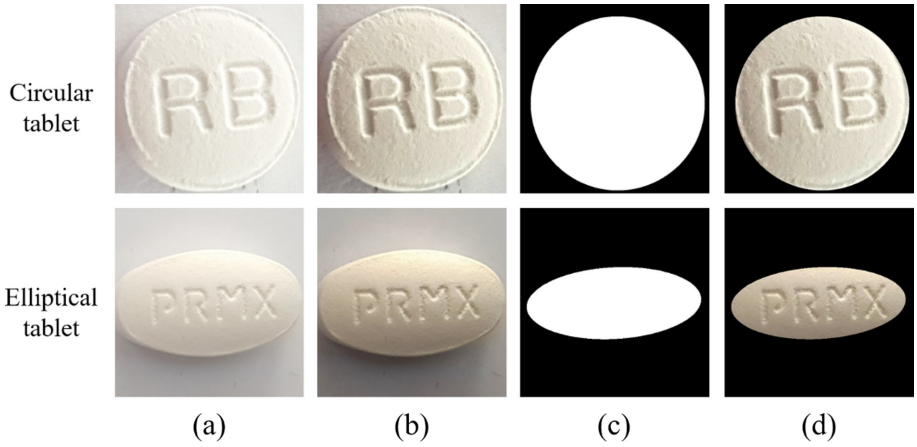

그림 2. 마스크 이미지와 원본 이미지의 합성 과정, (a) 원본 이미지, (b) 밝기 조절 이미지, (c) 마스크 이미지, (d) 합성된 이미지
Fig. 2. The process of merging the mask image and the raw image, (a) raw image, (b) brightness corrected image, (c) mask image, (d) merged image

미지는 그레이스케일(grayscale)로 변환 후 대비 제한 적응 히스토그램 평활화(CLAHE)[15]를 적용하였다.

이미지에서 글자 정보를 얻기 위해서는 글자 영역의 검출, 그리고 글자 인식이 필요하다. 본 논문에서는 영역 검출을 수행하는 CRAFT 모델[16]과 글자 인식을 수행하는 CRNN 모델[17]이 더해진 Keras-OCR을 사용하여 글자 정보를 문자열로 획득하였다. Keras-OCR은 그 성능이 검증된 OCR 오픈소스 API로, 빛 반사가 존재하는 이미지의 글자 인식에서 우수한 정확도를 보인 바 있다[18,19]. 이미지 1장 당 글자 정보는 분할선의 유무, 글자 배열 등에 따라 한 단어 또는 여러 단어로 인식되었고, 여러 단어로 인식된 경우 인식 순서대로 병합하여 최종적으로는 이미지당 하나의 단어를 출력하였다.

4. 알약 식별 모델

제안하는 모델은 이미지와 글자 정보 각각의 텐서(tensor)를 병합하여 알약의 카테고리를 분류하는 다중 입력 모델이고, 케라스를 이용한 함수형 모델(Functional API)[20,21] 형태를 구현하였다. 함수형 모델은 각 레이어를 일종의 함수로 정의하고 함수를 조합할 수 있는 연산자를 제공함으로써, 여러 레이어를 공유하거나 다양한 종류의 입력과 출력이 필요한 복잡한 모델의 설계를 가능하게 한다는 장점이 있다[22,23]. 본 논문에서 제안하는 모델의 구조는 Fig. 3과 같다.

CNN(Convolution Neural Network)의 구조는 기본적으로 Convolutional layer와 Pooling layer, 그리고 Fully-connected layer로 구성된다[24,25]. 제안 모델은 Convolutional 2D layer와 Pooling layer를 사용하여 이미지 정보를 획득하

www.kci.go.kr

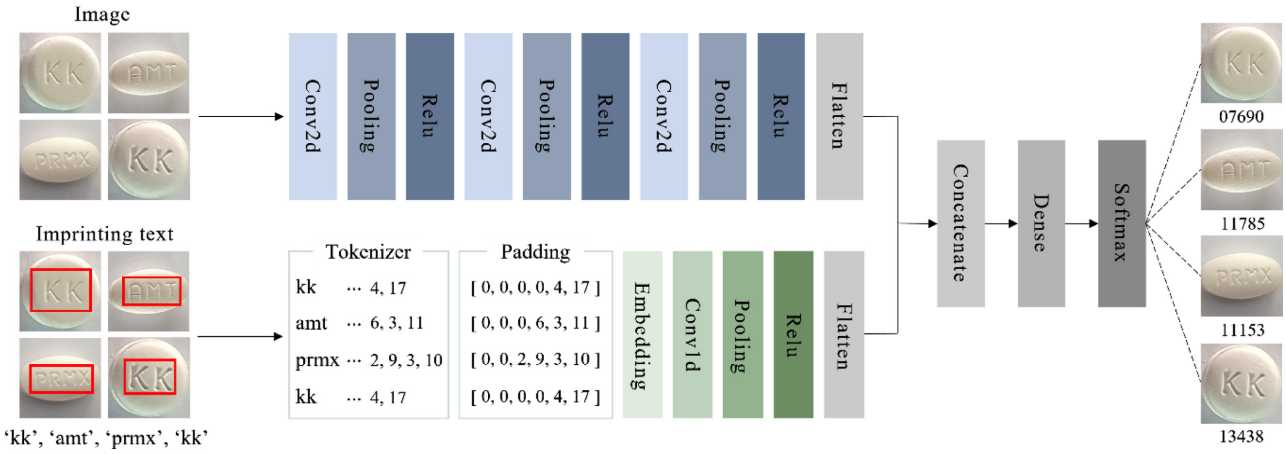

그림 3. 함수형 API에 기반한 제안 모델의 구조도
Fig. 3. Proposed architecture based on functional API

였다. 각 학습 이미지의 Ground truth인 문자는 Keras에서 제공하는 Tokenizer[26,27]를 이용하여 정수로 인코딩하고, 최대 단어 길이를 기준으로 Pre-padding[28]을 수행하였다. 이후 Embedding layer, Convolutional 1D layer와 Pooling layer를 사용하여 글자 정보를 획득하였다. 이미지와 글자 정보의 특징을 추출한 각각의 레이어는 concatenate 함수[29]를 이용하여 하나의 레이어로 병합하고, 이후 Fully connected layer와 softmax 함수[30]를 사용하여 알약 이미지를 식별하였다. 모델 학습에 사용된 파라미터 중 배치 사이즈(batch size)는 16, 반복 횟수는 50, 학습 조기종료(early stopping)의 함수인 patience는 20으로 설정하였다.

III. 결 과

본 논문에서는 음각 인식 모델과 알약 식별 모델을 각각 평가하였다. Keras OCR을 이용한 알약 음각 인식은 이미지 100장 중 90장의 단어를 정확하게 인식하여, 그 성능은 정확도 90%로 추정하였다. 알약의 음각을 제대로 인식하지 못한 10건 중, 1건은 문자열의 순서가 바뀌는 오류가 발생했고, 나머지 9건은 1글자씩 인식이 불가하거나 다른 문자로 인식하는 오류가 발생하였다.

1000장의 데이터는 10-fold 교차검증을 수행하여 학습 데이터까지 모두 한 번씩 검증되도록 하였고, 식별 모델의 검증 결과는 정확도(Accuracy), 정밀도(Precision), 재현율(Recall), F1-score를 사용하여 평가하였다[31]. 정확도와 정밀도, 재현율, F1-score는 오차행렬의 TP(True Positive), FP(False Positive), TN(True Negative), FN(False Negative)을 이용한 아래 식 (1)~(4)으로 계산하였다.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \tag{1}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{2}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{3}$$

$$\text{F1-score} = 2 \times \frac{(Precision \times Recall)}{Precision + Recall} \tag{4}$$

그 결과는 Table 1에 제시하였다. 제안하는 모델의 정확도 97.6%, 정밀도 97.8%, 재현율 97.6%, F1-score 97.5%를 확인하였고, 이미지를 단독 식별한 CNN 모델의 정확도 91.9%, 정밀도 93.1%, 재현율 91.9%, F1-score 91.8%를 확인하였다. One-vs-Rest 기법을 이용하여 ROC(Receiver operating

표 1. 딥러닝을 이용한 각 식별 모델의 성능 결과
Table 1. Result of each classification model based on deep learning

	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F1-score (95% CI)
Proposed classifier	0.973 (0.971-0.975)	0.974 (0.972-0.976)	0.973 (0.971-0.975)	0.972 (0.970-0.974)
2D CNN classifier	0.871 (0.868-0.874)	0.883 (0.880-0.886)	0.871 (0.868-0.874)	0.869 (0.866-0.872)
Statistical significance (p-value)	< 0.001	< 0.001	< 0.001	< 0.001

www.kci.go.kr

characteristic) Curve를 확인한 결과 두 모델의 AUC (Area under the curve)는 통계적으로 유의한 차이가 없었던 반면, 두 모델의 정확도, 정밀도, 재현율, F1-score는 모두 통계적으로 유의한 차이가 있음을 확인하였다(p<0.001).

IV. 고 찰

본 논문은 알약 이미지에서 음각 글자와 이미지의 정보를 각각 획득하여 함께 식별하는 모델을 제안하고, 이미지의 정보만을 획득하여 식별하는 모델과 그 성능을 비교하였다. 그 결과, 제안하는 모델의 정확도, 정밀도, 재현율, F1-score는 이미지 단독 식별 모델에 비해 통계적으로 유의한 우수성을 나타내었다.

글자 인식에 실패한 10장 중 7장이 최종 식별에 실패한 것을 보아, 인식된 글자가 정확하지 않을 경우 해당 이미지의 식별 정확도는 매우 떨어지는 것으로 판단된다. Fig. 4는 글자 인식에 실패한 데이터 일부를 나타내었다. 모양과 색상 정보가 유사한 알약은 각인된 글자로 구분되는데, 교차검증 결과 알약 100건 중 최종 식별율이 50% 미만인 2건은 글자가 같고 모양과 색상이 유사한 데이터가 존재하여 식별이 어려웠던 것으로 판단된다. 2건의 알약 이미지와 혼동된 알약들의 이미지는 각각 Fig. 5에 비교 제시하였다. 이러한 문제는 알약의 앞면과 뒷면 글자 모두와 분할선의 유무, 알약 실측 등 식별에 사용되는 정보의 범위를 확대하여 해결이 가능하다. 본 연구에서는 앞면 이미지만을 대상으로 연구를 수행했다는 제한점이 존재하므로, 추후 앞면과 뒷면의 이미지 정보를 각각 획득하고 글자 정보에 더하여 분류를 수행하는 연구를 진행하고자 한다.

본 연구에 사용된 테스트셋 데이터는 1클래스 당 데이터의 규모가 적고, 알약 회전 각도가 크지 않다는 한계가 있다. 이에 다양한 각도로 회전한 알약 이미지를 수집하고, 테스트 데이터의 비율을 확대하여 성능을 평가하는 후속 연구가 필요하다. 또한 글자를 인식할 때 이미지의 밝기를 일괄 조정하는 방법은 데이터 간 편차를 보정하기에 부족하다. 이는 이미지 각각의 밝기값을 고려하여 전처리를 수행하고, 알약의 표면 내에서 밝기 편차가 심할 경우 과도한 빛 반사 및 그림자를 제거하여 해결할 수 있다.

제안한 글자 인식 모델은 Suntronsuk의 결과에 따라 음

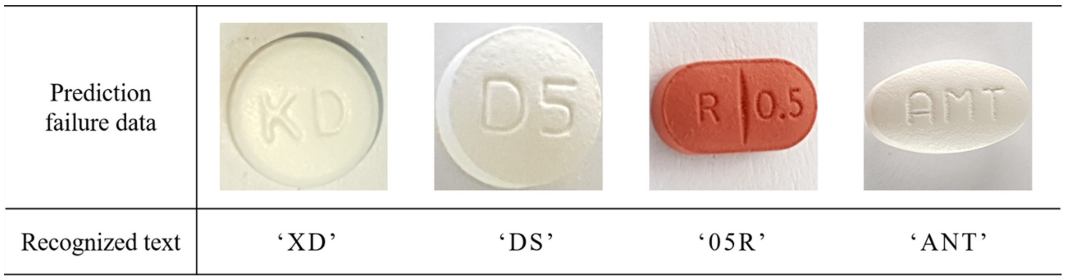

그림 4. 제안하는 모델의 식별 실패 데이터
Fig. 4. The result of prediction failure data for the proposed model

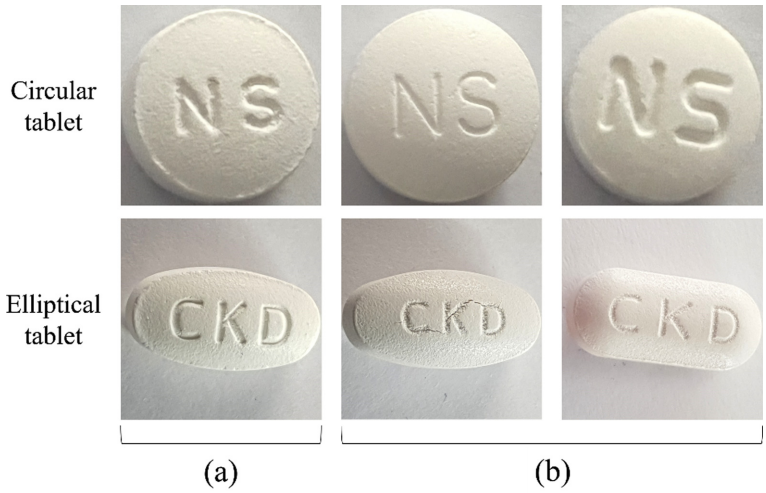

그림 5. 글자가 같은 하얀색 원형 알약들의 앞면 이미지, (a) 식별에 실패한 데이터, (b) 모양, 색깔, 글자 정보가 유사한 데이터
Fig. 5. The front side of white round tablets with the same text feature, (a) Prediction failure data, (b) The other data with similar shape, color, and text feature

www.kci.go.kr

각 글자의 인식률이 매우 낮다는 점을 고려하였고, 프린팅된 글자를 배제하여 음각 글자의 인식률 개선에 집중하였다. 음각 글자는 알약마다 음각의 면적, 깊이 등이 달라 글자의 인식률이 천차만별이다. 제시한 결과는 사용된 데이터에 한해 높은 인식률을 보였으나, 음각의 특징에 상관없이 일관된 성능을 보이기 위해서는 다양한 시도가 필요하다. 그 일례로, Keras OCR의 구조 변형을 시도할 수 있다. Keras OCR은 글자 영역 검출을 수행하는 CRAFT와 글자 인식을 수행하는 CRNN을 결합한 것으로 레이어, 파라미터 등의 변형이 가능하다. 구조 변형에 따른 성능을 평가한다면 음각 글자에 최적화된 모델 개발이 가능할 것이다.

일반적으로 이미지 식별 모델은 합성곱 레이어와 풀링 레이어를 반복 배치하고, 신경망이 깊어질수록 높은 성능을 보인다. 제안하는 식별 모델은 이미지 정보를 획득하는 신경망이 깊지 않고, 이는 최종 식별 성능과도 관계가 있을 것으로 판단된다. 추후 테스트 데이터의 비율이나 알약의 품목을 증대할 경우 심층 신경망의 설계가 필요할 것이다.

본 논문은 글자 정보를 인식할 때 알약마다 하나의 단어를 출력하고, 단어 자체를 학습 모델에 입력하여 정보를 획득하였다. 이는 글자 영역이 추출된 이미지를 입력한 선행연구에 비해 정보 획득이 더 수월하다는 장점이 있다. 한편 글자 인식만을 시도하거나, 인식한 글자 정보와 그 외 정보를 더한 분류 결과만을 도출한 선행연구와 달리, 본 논문은 알약 정보에 기반한 식별에서 글자 정보의 포함 여부로 나뉜 결과를 각각 제시 및 비교하였다. 이는 알약 식별에 있어서 글자 인식률 개선에 따른 효과를 검증하였다는 의의가 있다.

제언한 바를 토대로 글자 인식과 알약 식별에 다양한 방법을 시도하고, 그 결과 가장 적합한 모델을 개발한다면 현재 유통되는 국내 모든 알약의 식별이 가능하다. 후속 연구를 통해 더 많은 종류의 알약에서 정확한 식별 결과를 확인하고, 식별 모델이 적용된 알약 관련 어플리케이션을 구현할 경우 일반 사용자의 정보 접근성이 높아질 것이다. 더 나아가 식별 모델이 응용된 알약 분류기는 조제 오류로 발생하는 약화사고의 발생 가능성을 미연에 방지할 수 있고, 이러한 순기능이 알약 자동 식별 모델의 상용화를 촉진할 것으로 기대한다.

Reference

[1] Oh DE, Lee MS, Lee YD, Ye KN, and Kim JT. *The reduction of unreturned medication by improving the various medication return activities*. Journal of Korean Society of Health-system Pharmacists. 2011;28(4):364-371.

[2] Eoum GH, Lee JY, Cho YH, Cho YS, Hahn HJ, and Son IJ. *The Effect for Pharmacy Intervention on the Decrease of Returned Medications from the Ward*. Journal of Korean Society of Health-system Pharmacists. 2006;23(3):1-9.

[3] Yi GY, Kim YJ, Kim ST, Kim HE, and Kim KG. *Comparison and Verification of Deep Learning Models for Automatic Recognition of Pills*. Journal of Korea Multimedia Society. 2019;22(3):349-356.

[4] Wang Y, Ribera J, Liu C, Yarlagadda S, Zhu F. *Pill Recognition Using Minimal Labeled Data*. 2017 IEEE Third International Conference on Multimedia Big Data (BigMM). 2017; 346-53.

[5] Larios Delgado N, Usuyama N, Hall AK, Hazen RJ, Ma M, Sahu S, et al. *Fast and accurate medication identification*. NPJ digital medicine. 2019;2(1):1-9.

[6] Ou YY, Tsai AC, Zhou XP, Wang JF. *Automatic drug pills detection based on enhanced feature pyramid network and convolution neural networks*. IET Computer Vision. 2020; 14(1):9-17.

[7] Suntronsuk S, Ratanotayanon S, editors. *Automatic text imprint analysis from pill images*. 2017 9th International Conference on Knowledge and Smart Technology (KST). 2017; 288-293.

[8] Kim DW. *Shape and Text Imprint Recognition of Pill Image Taken with a Smartphone*. Master's Thesis of Seoul National University. 2017.

[9] Ko DG. *Optical Character Recognition Performance Comparison of CNNs and Tesseract*. Master's Thesis of Sungkyunkwan University, 2016. [Online]. Available: http://www.riss.kr/link?id=T14177015

[10] Kim YJ and Kim KG. *Development of an Optimized Deep Learning Model for Medical Imaging*. Journal of the Korean Society of Radiology. 2020;81(6):1274-1289.

[11] Hassanein A.S, Mohammad S, Sameer M, Ragab M.E. *A survey on Hough transform, theory, techniques and applications*. International Journal of Computer Science Issues (IJCSI). 2015;12(1):139-156.

[12] Xu Z, Baojie X, Guoxin W. *Canny edge detection based on Open CV*. 2017 13th IEEE international conference on electronic measurement & instruments (ICEMI), 2017;53-56.

[13] Firoz R, Ali M.S, Khan M.N.U, Hossain M.K, Islam M.K, Shahinuzzaman M. *Medical image enhancement using morphological transformation*. Journal of Data Analysis and Information Processing. 2016;4(1):1-12.

[14] Erkan U, Gökrem L, Enginoğlu S. *Different applied median filter in salt and pepper noise*. Computers & Electrical Engineering, 2018;70:789-798.

[15] Sahu S, Singh A.K, Ghrera S, Elhoseny M. *An approach for de-noising and contrast enhancement of retinal fundus image using CLAHE*. Optics & Laser Technology, 2019;110:87-98.

[16] Baek Y, Lee B, Han D, Yun S, Lee H. *Character region awareness for text detection*. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019; 9365-9374.

[17] Lei Z, Zhao S, Song H, Shen J. *Scene text recognition using residual convolutional recurrent neural network.* Machine Vision and Applications. 2018;29(5):861-871.

[18] Chotivatunyu P, Hnoohom N. *Medicine Identification System on Mobile Devices for the Elderly*. 2020 15th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP). 2020;1-6.

[19] Maitrichit N, Hnoohom N. *Intelligent Medicine Identification System Using a Combination of Image Recognition and Optical Character Recognition*. 2020 15th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP). 2020;1-5.

[20] Manaswi NK. *Understanding and working with Keras*. Deep

www.kci.go.kr

Learning with Applications Using Python. Berkeley, CA: Apress: 2018;31-43.

[21] Chicho BT, Sallow AB. *A Comprehensive Survey of Deep Learning Models Based on Keras Framework*. Journal of Soft Computing and Data Mining. 2021;2(2):49-62.

[22] Chollet F, *Deep learning with Python*. New York: Simon and Schuster: 2021;62-64.

[23] Gulli A, Kapoor A, Pal S. *Deep learning with TensorFlow 2 and Keras: regression, ConvNets, GANs, RNNs, NLP, and more with TensorFlow 2 and the Keras API*. Birmingham, UK: Packt Publishing Ltd: 2019;63-66.

[24] Albawi S, Mohammed TA, Al-Zawi S. *Understanding of a convolutional neural network*. 2017 international conference on engineering and technology (ICET). 2017;1-6.

[25] Sun M, Song Z, Jiang X, Pan J, Pang Y. *Learning pooling for convolutional neural network*. Neurocomputing. 2017;224:96-104.

[26] Vijayarani S, Janani R. *Text mining: open source tokenization tools-an analysis*. Advanced Computational Intelligence: An International Journal (ACII), 2016;3(1):37-47.

[27] Kudo T, Richardson J. *Sentencepiece: A simple and language independent subword tokenizer and detokenizer for neural text processing*. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018;66-71.

[28] Dwarampudi M, Reddy N. *Effects of padding on LSTMs and CNNs*. arXiv preprint arXiv:1903.07288, 2019.

[29] Du C, Wang Y, Wang C, Shi C, Xiao B. *Selective feature connection mechanism: Concatenating multi-layer CNN features with a feature selector*. Pattern Recognition Letters. 2020;129:108-14.

[30] Ertam F, Aydın G. *Data classification with deep learning using Tensorflow*. 2017 international conference on computer science and engineering (UBMK), 2017;755-758.

[31] Yacouby R, Axman D. *Probabilistic extension of precision, recall, and F1 score for more thorough evaluation of classification models*. Proceedings of the first workshop on evaluation and comparison of NLP systems. 2020;79-91.

www.kci.go.kr