

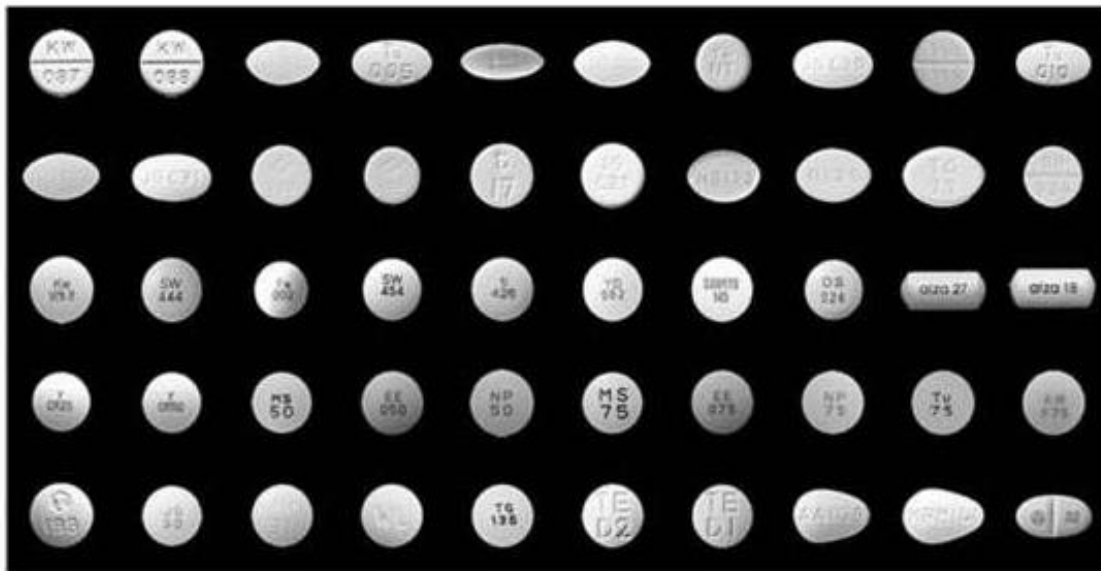
스마트폰으로 촬영된 알약 영상의 글자 및 형상 인식 방법

<http://s-space.snu.ac.kr/bitstream/10371/137361/1/000000146053.pdf>

- 기존의 알약 detection 논문의 한계

1. 모양과 색상 일부를 사용하여 detect하는 알고리즘들은 광범위한 알약 종류들에 대해서는 적용하기 힘들다.
2. 사용한 이미지가 대부분 오른쪽 사진과 같은 **제한된 환경 내의 이미지**

그래서, 실제 사용자가 찍은 이미지로는 낮은 인식률



제한된 환경내의 이미지



실제 사용자가 촬영한 알약 이미지

- 본 논문에서 제안하는 것

- : 스마트폰으로 촬영된 알약 영상에서 글자와 형상 인식

- : 알약 영역을 특정하기 위해 Saliency Map 이용한 뒤, 빛/그림자 제거 → 알약 영역 획득

- : 위에서 획득한 영역에서 Zernike Moment를 통해 형상정보를 얻음

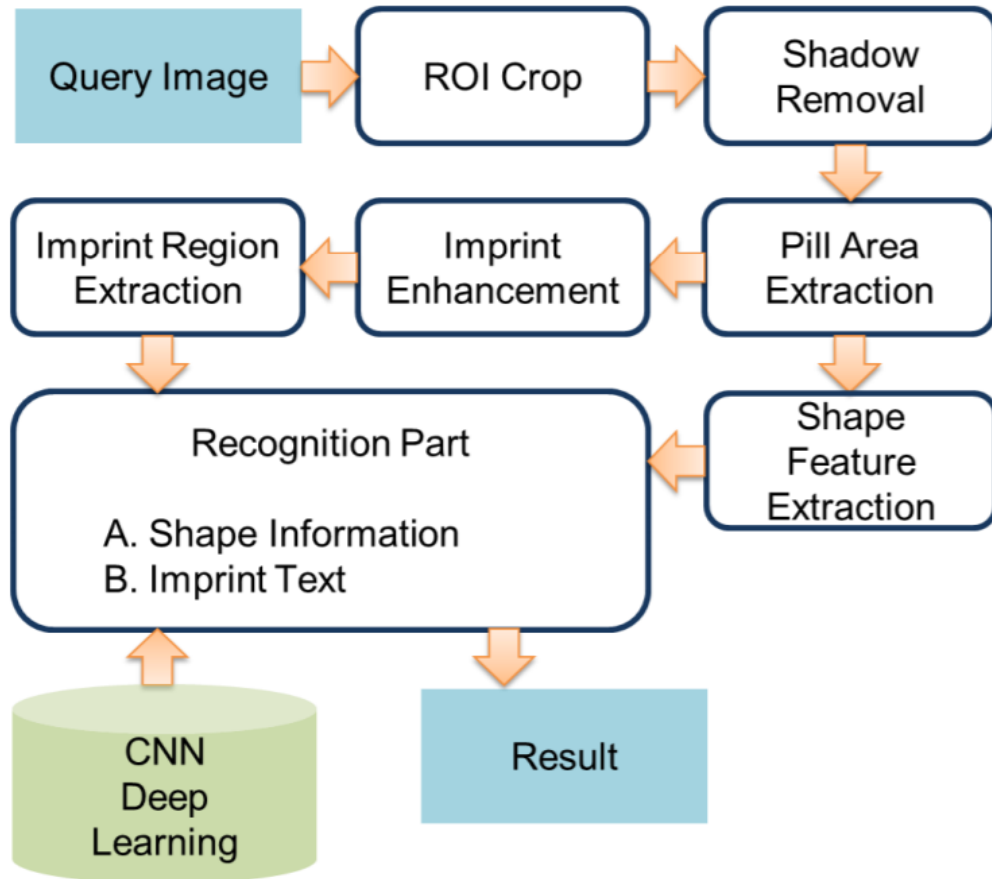
- : Gaussian Filter, Gradient Filter, Binarization을 통해 글자박스 추출 → CNN을 통해 OCR

- 전체 시스템 개요

- : 알약 주변부 획득 → 그림자/빛 효과 제거 → 배경과의 색상 차이 및 Canny Edge를 이용하여 알약 영역을 추출

- Zernike Moment를 이용해 형상 정보 획득 → Gradient filter를 이용해 글자부를 강조 → 글자의

- Bounding Box 획득 → CNN을 통해 글자 읽음 → 모양 정보 + 글자 정보 종합하여 알약 인식



Bounding Box를 획득 한 후

CNN을 통해 각각의 글자부를 읽음으로써 글자 정보 획득
→ 최종적으로 모양 정보와 글자 정보를 종합하여 알약 인식

- CNN Deep learning : 글자 학습을 따로 하여 각인 글자에 특화된 OCR

- 사용한 데이터

: NLM DB의 Reference Image에서 글자들을 수동 추출하고 합성 이미지를 만듦

: 총 36개의 글자

: 글자 별 학습에 사용한 데이터 수 → 수동 추출 이미지 2000장 + 합성 이미지 1000장 = 총 3000장



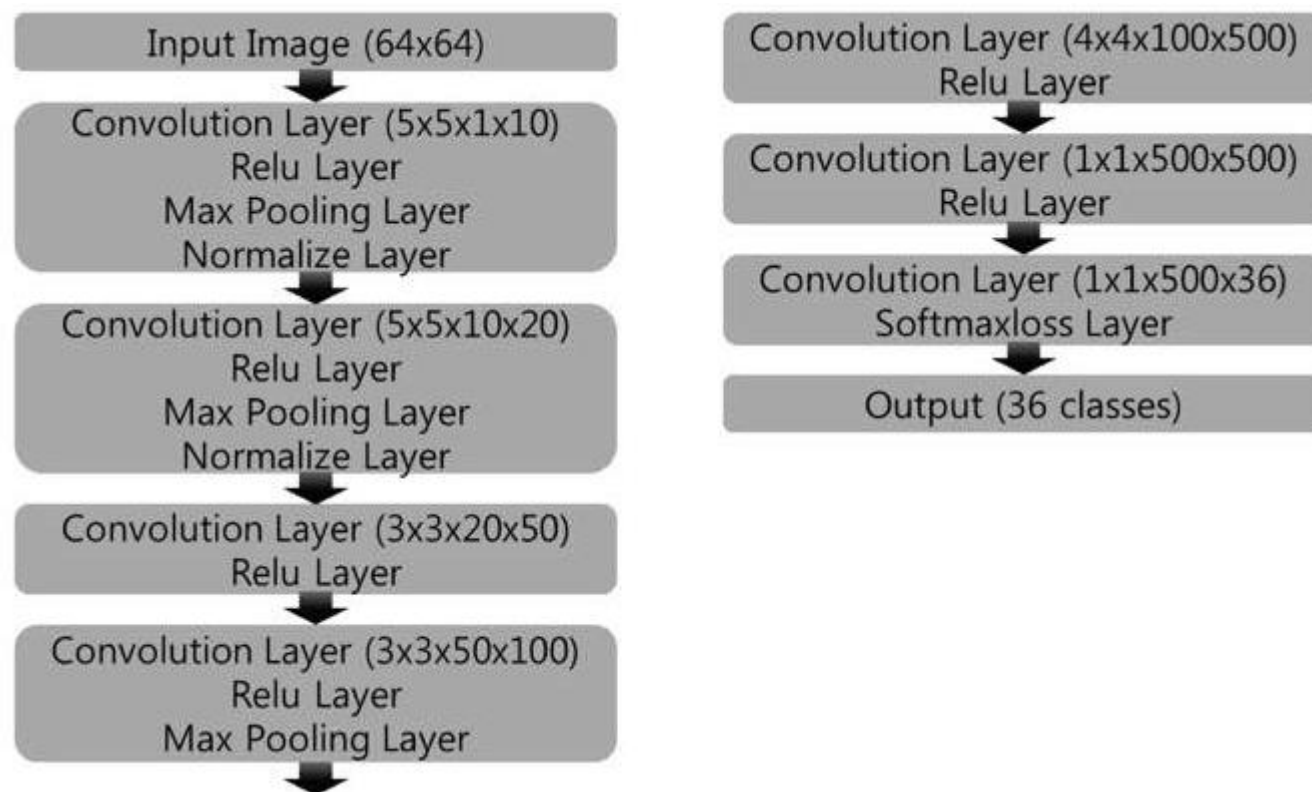
Consumer Quality Image

Reference Image

→ NLM 이미지 DB

- CNN Deep learning : 글자 학습을 따로 하여 각인 글자에 특화된 OCR

- 네트워크 구성



- 결과

- Bounding Box detection Rate

기존의 연구 →

	Proposed	Suntronsuk et al. (2016)	Kasar et al. (2007)
Text Box Detection Rate	87.5%	56.67%	21.21%

- Character Recognition Rate

: Box detection과는 별개로,
알약에 존재하는 글자들을 얼마나 찾았는가/
얼마나 정확하게 찾았는가

	Proposed	Suntronsuk et al. (2016)	Kasar et al. (2007)
Precision	0.776	0.388	0.144
Recall	0.835	0.410	0.152
F-score	0.786	0.399	0.148

- 결과

- Text Recognition Rate & Recognition Success Rate

: 글자(character)들이 모여서 하나의 Text를 이루고, Text가 알약을 판별하는 구분기호로 사용됨

		Consumer Quality Image		Reference Image	
		Proposed	Suntronsuk et al. (2016)	Proposed	Jiye Yu et al. (2015)
Text Recognition Rate	Without parsing	55.4%	22.61%	73.2%	-
	With parsing	73%	-	92.8%	-
Recognition Success rate (with parsing)		58.4%	-	90.4%	90.46%

↓
글자와 모양을 모두 포함한 인식률

- 단순한 Text 인식
: 55.4%
- Parsing을 통해 DB에 존재하지 않는 경우 가장 연관성이 높은 다른 Text로 검색하여 대체한 경우
: 73%
: 나머지 27%의 오류
→ 글자가 한 글자인 경우 多
(글자가 한글자인 경우 Box가 크게 잡히게 되고, 경우에 따라 이를 작은 단위로 나누기 때문)

- 인식률을 높이기 위해서는

- 글자 크기에 영향을 덜 받는 알고리즘을 구축

- 글자 인식에 실패한 경우, 사용자에게 재촬영 요구

- 모양 인식률과 글자 박스 추출 성능을 개선한다면,

인식률은 더욱 개선될 것이며 일반 사용자들에게도 충분히 적용가능