

6주차 Mini project – 텍스트 인식

강용진

```
lookup.KeyValue  
f.constant(['em  
=tf.constant([0  
ce = tf.lookup.StaticV  
init,  
num_oov_buckets=5)
```

```
lookup.StaticVocabular  
initializer,  
num_oov_buckets,  
lookup_key_dtype=None  
name=None,  
experimental_is_open
```

알약 텍스트 인식의 문제점

텍스트 형태



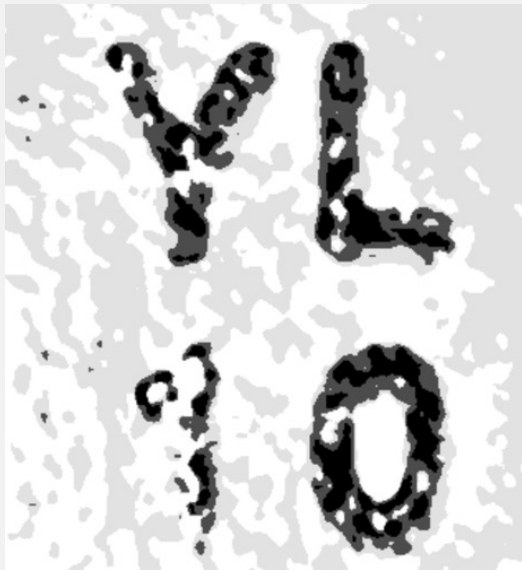
- 일반적인 형태가 아닌 경우

표면 질감



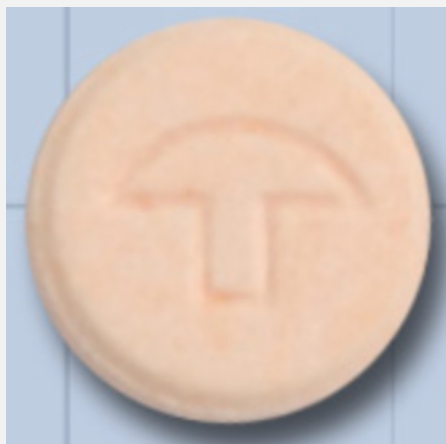
- 이진화, 흑백 처리를 거쳤을 때 울퉁불퉁한 부분이 글자와 구분되지 않음
- 나정, 필름코팅정 등 여러 제형에서 나타남

질감에 기인하는 오인식



- 질감이 텍스트의 형태에 영향을 미쳐 오인식을 유발하기 쉬움 (위 사진: Y를 ¥라고 인식)

기타 해결이 어려운 문제점



- 특이 폰트, 텍스트가 아닌 경우 등

해결방안, 목표

해결방안, 목표

해결하는 방향

: 질감, 형태 등은 경우에 맞는 전처리 방법을 모색
전처리 기준 찾기(제형에 따라? 색 분포에 따라?)

텍스트 정보를 사용하지 않는 방향

: 확실하게 인식이 가능한 경우가 아니라면 텍스트 정보를 고려하지 않는 것이 나음