



...

3팀의 여행가이드

...

존 경 받 는 조 장 팀

임 준 규 · 김 태 운 · 류 덕 현 · 이 인 아 · 조 민 경



INDEX



- 주제
- 키워드 추천
- 이미지 분석
- 결과



주제



존 경 받 는 조 장 팀



주제 선정

“일본 **교토**를 가는 여행객에게 도움이 되는 서비스를 만들자”



기능

1. 키워드에 해당하는 맛집 추천
2. 이미지 인식으로 정보 제공



키워드 추천



존 경 받 는 조 장 팀

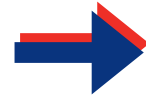




키워드 추천 - 데이터

존 경 받 는 조 장 팀

Google



食べログ

食べログ

맛있는 레스토랑을 찾아 보자!!

NAVER



키워드 추천 - 타베로그

존 경 받 는 조 장 팀



<홈페이지>

Unnamed: 0	가게이름	블로그리뷰
0	銀座ライオン 京都四條烏丸店	Vlnlrn この日のランチタイムは、Aさんと阪急烏丸駅直結のビルの地...
1	銀座ライオン 京都四條烏丸店	Vlnlrn 銀座ライオンの京都四條烏丸店です。この日は、平日の夕刻ご...
2	銀座ライオン 京都四條烏丸店	Vlnlrn この日も烏丸で仕事やったんでランチは久しぶりにこちらへ！...
3	銀座ライオン 京都四條烏丸店	Vlnlrn こういうビヤレストラン、男性には絶大な人気。会社から行く...
4	銀座ライオン 京都四條烏丸店	Vlnlrn 上司に連れられての、何度か目の訪問。上司(50代前後)は...
5	銀座ライオン 京都四條烏丸店	Vlnlrn 大体土曜日のお昼に利用しています。ランチは日替わりのシェ...
6	銀座ライオン 京都四條烏丸店	Vlnlrn 昼からビールを飲んで、語らうのであれば良い感じの店。個室...
7	銀座ライオン 京都四條烏丸店	Vlnlrn 久しぶりに会社で銀座ライオンに、久々に行くこのビヤホー...
8	銀座ライオン 京都四條烏丸店	Vlnlrn 161027美味しいビールを飲みたい！取りあえず近場で...
9	銀座ライオン 京都四條烏丸店	Vlnlrn ちょっと四條烏丸で茶をしばきにと、候補に挙がったのがこちら。ただ、相方曰...
10	銀座ライオン 京都四條烏丸店	Vlnlrn ビールの飲み放題はやはりビアホールだけあって美味しく頂き...
11	銀座ライオン 京都四條烏丸店	Vlnlrn 説明不要なほど、全国にチェーン展開している、安定した品質...
12	銀座ライオン 京都四條烏丸店	Vlnlrn 爽いビールが美味しいお店ですね私が当日頂いたのは琥珀、ハー...
13	銀座ライオン 京都四條烏丸店	Vlnlrn ふらっと入りましたさすがに老舗というか昔からあるビヤレス...
14	銀座ライオン 京都四條烏丸店	Vlnlrn 京都四條烏丸、仕事終了後、地下鉄で、京都駅へ向かうシチュ...
15	銀座ライオン 京都四條烏丸店	Vlnlrn 四條烏丸のビヤレストランライオンに同僚と行きました。ここ...
16	銀座ライオン 京都四條烏丸店	Vlnlrn ランチ利用で月1回位利用しています。まあ、近いし使い勝手...
17	銀座ライオン 京都四條烏丸店	Vlnlrn 夜に伺いました。ビアホールですが、非常に落ち着いた雰囲気...
18	銀座ライオン 京都四條烏丸店	Vlnlrn "エビス・スタウト・クリーミートップ"を飲みに！かのギネ...
19	銀座ライオン 京都四條烏丸店	Vlnlrn ビールも グラスの状態も 恵比寿や札幌にくらべるとイマイ...
20	無門 三条室町店	Vlnlrn 母上様のダイエット進行中！！まだまだ道のりは続きますが第...
21	無門 三条室町店	Vlnlrn グルーボンりたいなチケットで、しゃぶしゃぶコースにドリン...
22	無門 三条室町店	Vlnlrn ランチタイムに伺いました。11:20頃についたらまだ準備...
23	無門 三条室町店	Vlnlrn 会社の宴会で利用しました。ここはお鍋が美味しいと聞いてま...

<리뷰 데이터프레임>



키워드 추천 - 동작과정

준 경 받 는 조 장 팀

✓ 키워드 증식

```
: model2 = Word2Vec.load('korean_word2vec.model')
```

```
C:\Users\acorn\Anaconda3\envs\py36\lib\site-packages\smart_open\smart_open_lib.py:398: UserWarning: T
details: https://github.com/RaRe-Technologies/smart_open/blob/master/README.rst#migrating-to-the-new-
'See the migration notes for details: %s' % _MIGRATION_NOTES_URL
```

```
: a = model2.wv.most_similar('세련되다')
a
```

```
: [('모던', 0.7309333086013794),
 ('인테리어', 0.7253825664520264),
 ('감각', 0.6635186672210693),
 ('단정하다', 0.6612066626548767),
 ('중후', 0.6115866899490356),
 ('꾸미다', 0.6090570688247681),
 ('고급스럽다', 0.6054890155792236),
 ('심플', 0.6037949919700623),
 ('고풍', 0.6017404794692993),
 ('휴게', 0.5840120315551758)]
```




키워드 추천 - 동작과정

존 경 받 는 조 장 팀

✓ 유사도 측정

```
1 translate = translation('분위기 좋은 와인바')
2 a = model_jp.infer_vector(parse_text(translate).orig)
3 r1 = model_jp.docvecs.most_similar([a], topn=20)
4 r1
```

```
Ⓜ /usr/local/lib/python3.6/dist-packages/gensim/matutils.py:737: FutureWarning
  if np.issubdtype(vec.dtype, np.int):
[('트렌타노베', 0.5513025522232056),
 ('로비어 라운지', 0.5325621962547302),
 ('로비어 라운지', 0.529084801673889),
 ('L'ARTS', 0.5247204900707888),
 ('メルキュール カフェ', 0.5225375890731812),
 ('CAVA', 0.5202914476394653),
 ('妃屋', 0.5157334804534912),
 ('イタリア食堂 3000ガロッパレ', 0.5102201104164124),
 ('piu', 0.5075822472572327),
 ('ラテアート・ジャンキーズ・ロースティングショップ 北野天満宮店', 0.5067816),
 ('HOTOKI', 0.5014253854751587),
 ('アマコリス', 0.5001779198646545),
 ('ピッツァ フォールム', 0.49935266375541687),
 ('SOI F. CAFF', 0.4982700049877167).
```

결과





키워드 추천 - 동작원리

존 경 받 는 조 장 팀

✓ Word2Vec

중심 단어 주변 단어

↓ ↓

The fat cat sat on the mat

The fat cat sat on the mat

The fat cat sat on the mat

The fat cat sat on the mat

The fat cat sat on the mat

The fat cat sat on the mat

The fat cat sat on the mat

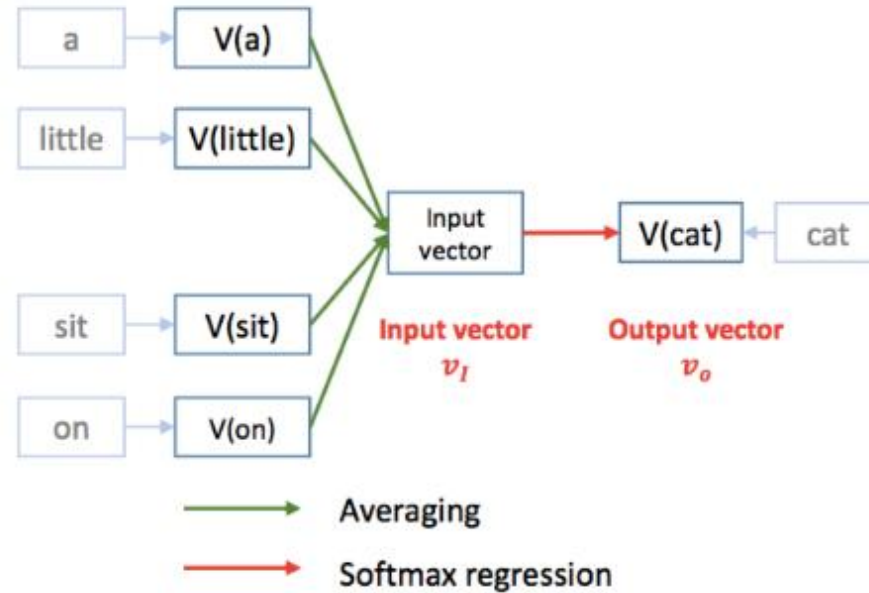
중심 단어	주변 단어
[1, 0, 0, 0, 0, 0, 0]	[0, 1, 0, 0, 0, 0, 0], [0, 0, 1, 0, 0, 0, 0]
[0, 1, 0, 0, 0, 0, 0]	[1, 0, 0, 0, 0, 0, 0], [0, 0, 1, 0, 0, 0, 0], [0, 0, 0, 1, 0, 0, 0]
[0, 0, 1, 0, 0, 0, 0]	[1, 0, 0, 0, 0, 0, 0], [0, 1, 0, 0, 0, 0, 0], [0, 0, 0, 1, 0, 0, 0], [0, 0, 0, 0, 1, 0, 0]
[0, 0, 0, 1, 0, 0, 0]	[0, 1, 0, 0, 0, 0, 0], [0, 0, 1, 0, 0, 0, 0], [0, 0, 0, 0, 1, 0, 0], [0, 0, 0, 0, 0, 1, 0]
[0, 0, 0, 0, 1, 0, 0]	[0, 0, 1, 0, 0, 0, 0], [0, 0, 0, 1, 0, 0, 0], [0, 0, 0, 0, 0, 1, 0], [0, 0, 0, 0, 0, 0, 1]
[0, 0, 0, 0, 0, 1, 0]	[0, 0, 0, 1, 0, 0, 0], [0, 0, 0, 0, 1, 0, 0], [0, 0, 0, 0, 0, 0, 1]
[0, 0, 0, 0, 0, 0, 1]	[0, 0, 0, 0, 1, 0, 0], [0, 0, 0, 0, 0, 1, 0]



키워드 추천 - 동작원리

존 경 받 는 조 장 팀

✓ Word2Vec

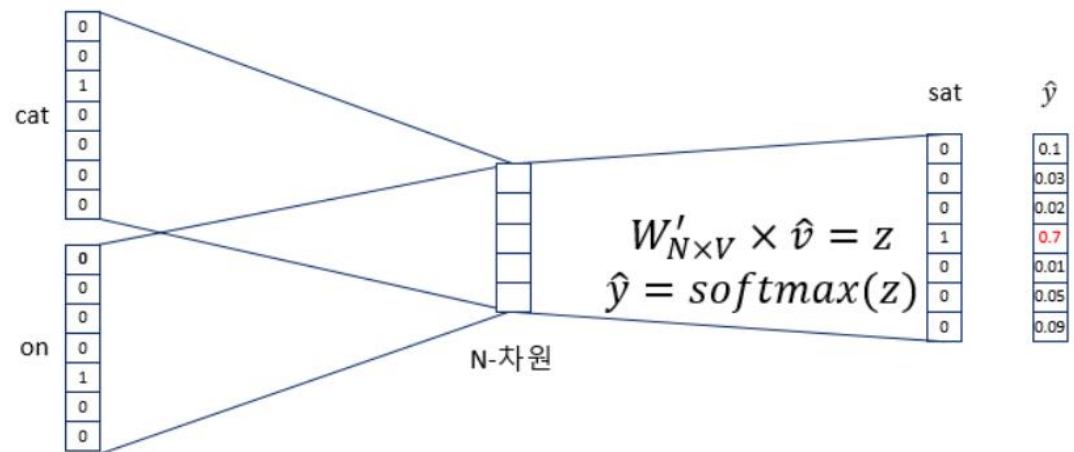
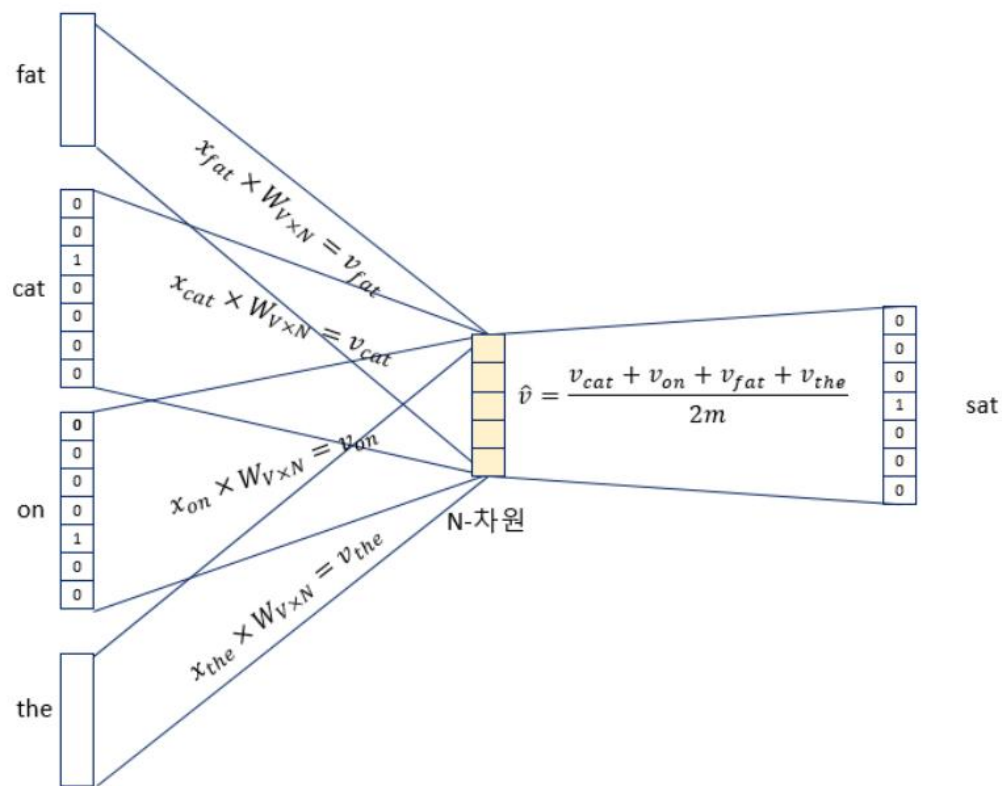




키워드 추천 - 동작원리

준 경 받 는 조 장 팀

✓ Word2Vec

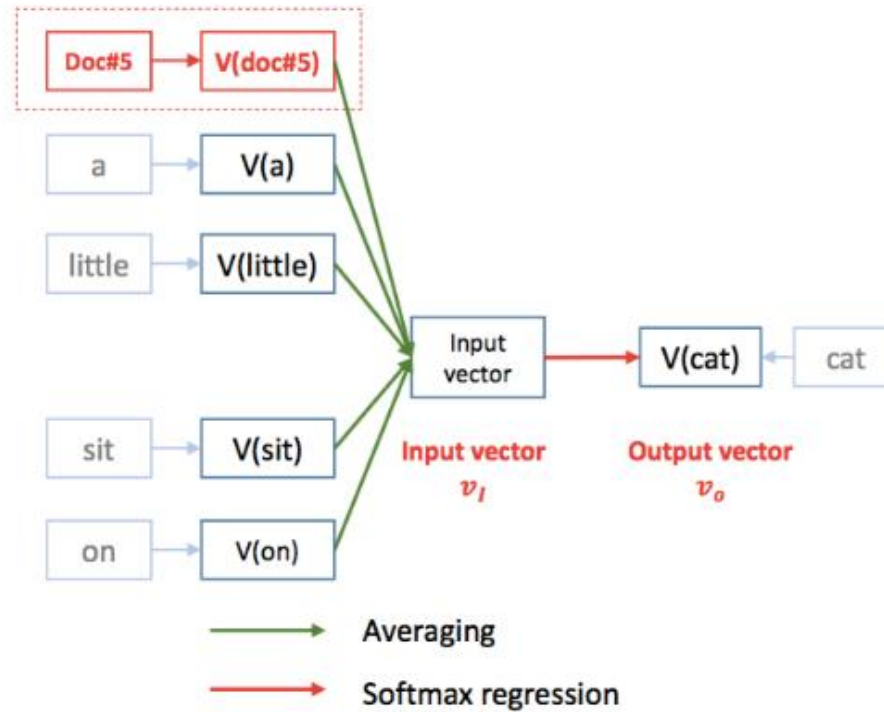




키워드 추천 - 동작원리

존 경 받 는 조 장 팀

✓ Doc2Vec





키워드 추천 - 보완점

존 경 받 는 조 장 팀

- ✓ TF-IDF
- ✓ BM25

$$w_{x,y} = \text{tf}_{x,y} \times \log\left(\frac{N}{\text{df}_x}\right)$$

TF-IDF

Term x within document y

$\text{tf}_{x,y}$ = frequency of x in y

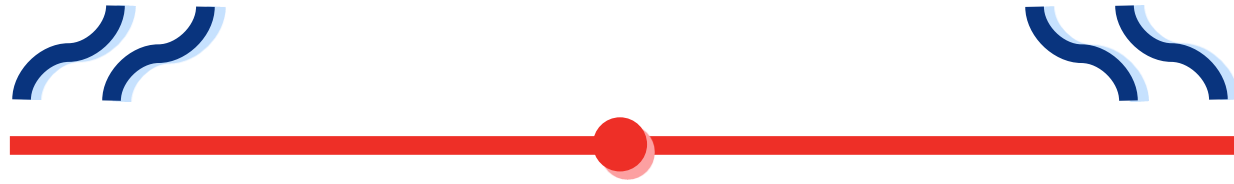
df_x = number of documents containing x

N = total number of documents

TF-IDF

$$\text{score}(D, Q) = \sum_{i=1}^n \text{IDF}(q_i) \cdot \frac{f(q_i, D) \cdot (k_1 + 1)}{f(q_i, D) + k_1 \cdot \left(1 - b + b \cdot \frac{|D|}{\text{avgl}}\right)},$$

BM25



이미지 분석



준 경 받 는 조 장 팀



건물 사진 분석

존 경 받 는 조 장 팀



37.499401, 127.029004

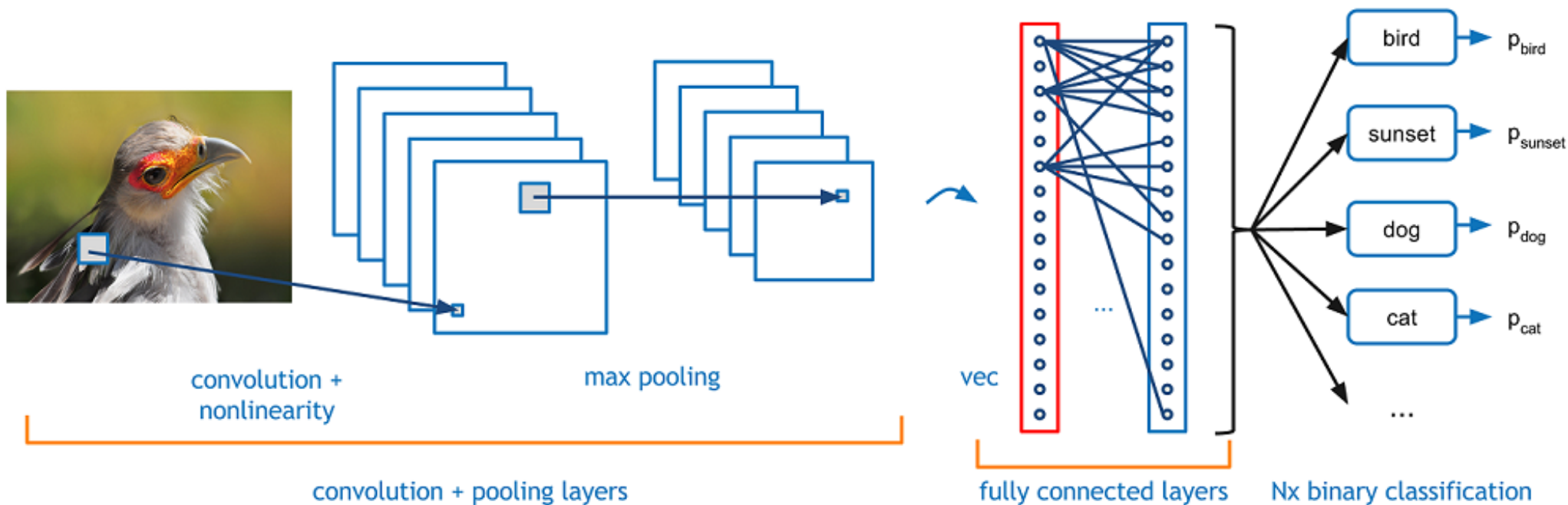
- [에이콘](#)
- [서울특별시 강남구 강남대로94길 20](#)
- [info](#)
- [etc](#)

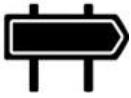
Close



건물 사진 분석 - CNN

존 경 받 는 조 장 팀





간판 사진 분석 - CRNN

존 경 받 는 조 장 팀

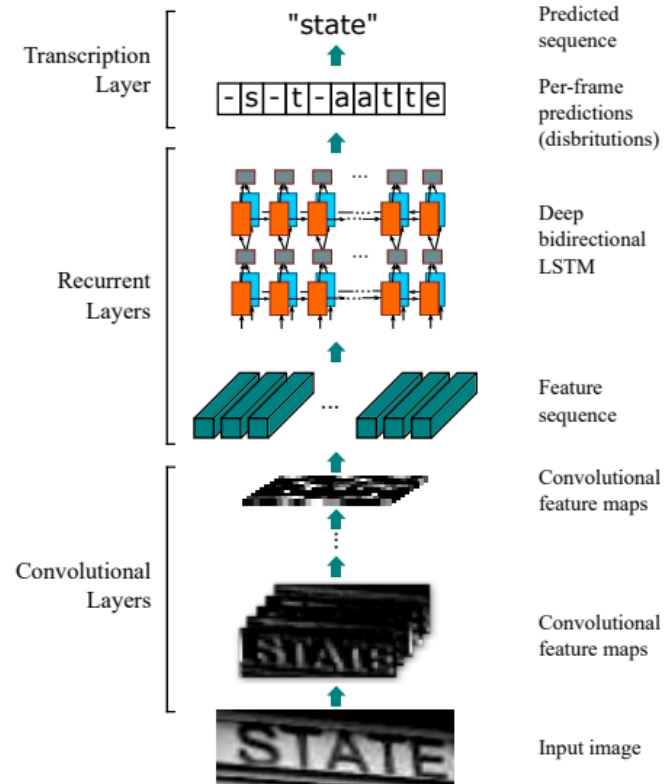


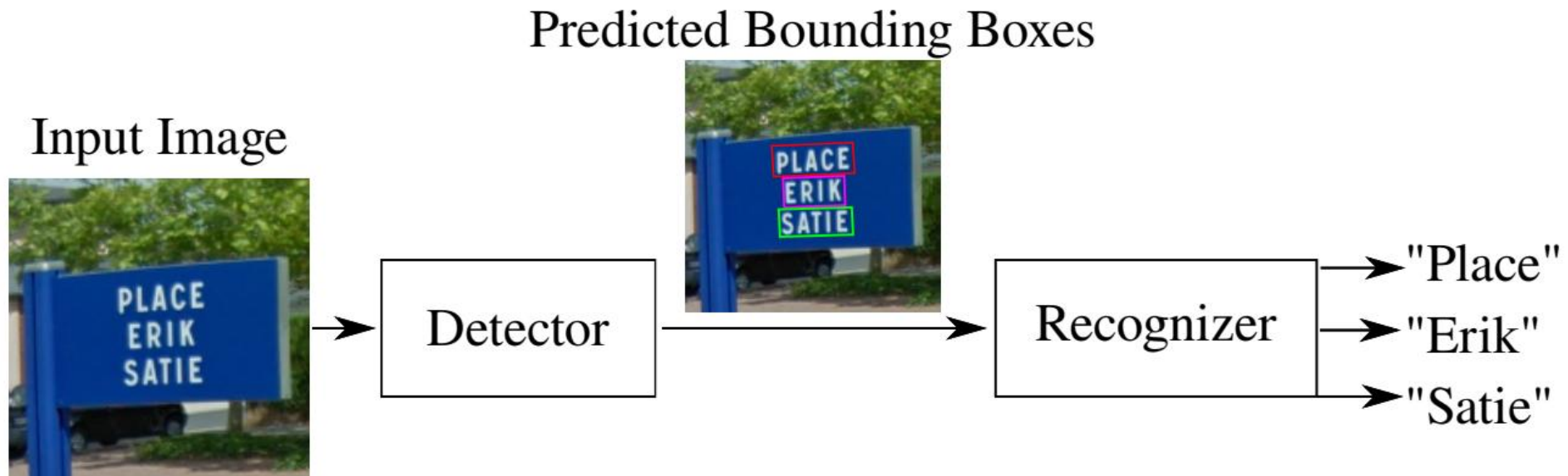
Figure 1. The network architecture. The architecture consists of three parts: 1) convolutional layers, which extract a feature sequence from the input image; 2) recurrent layers, which predict a label distribution for each frame; 3) transcription layer, which translates the per-frame predictions into the final label sequence.



간판 사진 분석 - CRNN

존 경 받 는 조 장 팀

✓ End to end





간판 사진 분석 - CRNN

존 경 받 는 조 장 팀

✓ 데이터셋 만들기

おせはる

u304a u305c
u306f u3091
u3050
_190716_225...

おせじをとら

u304a u305c
u3080 u3092
u3068 u3046
u3042 _1907...

おせんぬえ

u304a u305c
u3093 u306c
u3048
_190716_222...

おぞおそ

u304a u305e
u304a u305d
_190716_22203
0.png

おぞせはぶ

u304a u305e
u305b u306f
u3076
_190716_222...

おぞてにえね

u304a u305e
u3066 u306b
u3048 u306d
_190716_222...

おぞでぎめ

u304a u305e
u3067 u304e
u3081 u306c
_190716_225...

おぞへげん

u304a u305e
u3078 u3052
u3079
_190716_222...

おたく

u304a u305f
u304f
_190716_22203
0.png

おたしこなし

u304a u305f
u3057 u3053
u306a u3057
_190716_225...

おたあすのべり

u304a u305f
u3091 u305a
u306e u307a
u308a _19071...

おなすてへ

u304a u306a
u3059 u3067
u3078
_190716_225...

おにり

u304a u306b
u308a
_190716_22203
0.png

おにれたゆふび

u304a u306b
u308c u305f
u3086 u3076
u3074 _1907...

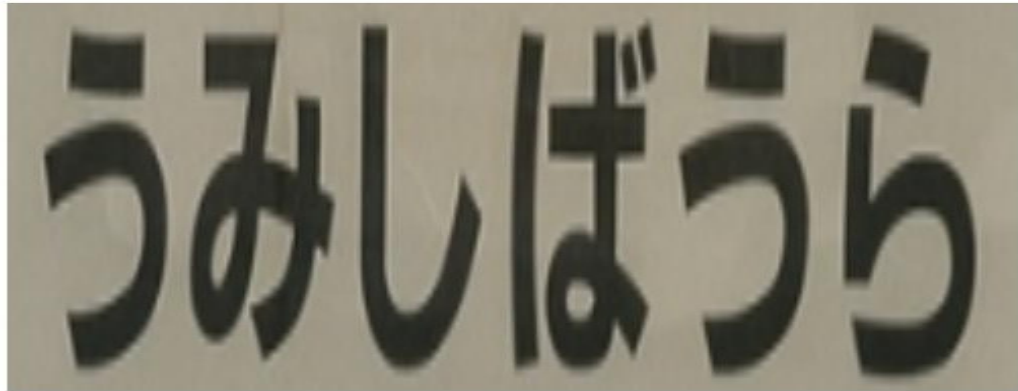
おにくじらゆやへ

u304a u306b
u3050 u3058
u3089 u3086
u3084 u3079 ...



간판 사진 분석 - 결과

존 경 받 는 조 장 팀



결과



うみしばうる



결과

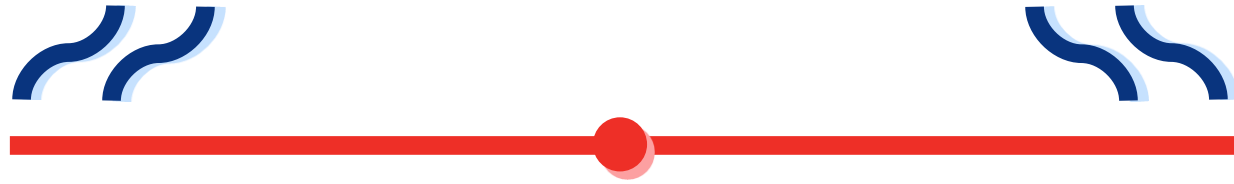


준 경 받 는 조 장 팀

결과

존 경 받 는 조 장 팀

<http://192.168.0.38:8080/>



Q & A





감사합니다!

존 경 받 는 조 장 팀

임 준 규 · 김 태 운 · 류 덕 현 · 이 인 아 · 조 민 경