

## POI(Point Of Interest) 이름 매칭 방법에 관한 연구

임재호, 최승민, 이재영  
한국전자통신연구원

### A study on POI(Point Of Interest) name matching technique

Jaeho Lim, Seung-Min Choi, Jae-Yeong Lee  
Electronics and Telecommunications Research Institute

**Abstract** - 이동체의 위치를 추정하기 위해서는 GPS 정보 뿐만 아니라, 상호보완적으로 사용할 수 있는 추가적인 정보들을 이용하는 것이 필요하다. 이러한 추가적인 정보들 중의 하나로 POI 정보를 이용할 수 있지만 OCR 수행결과가 완전하지 않는 경우가 많아서 이를 개선하기 위한 방법이 필요하다. 본 논문에서는 기존에 알고 있는 POI들의 이름에 OCR 수행결과로 오인식 되어 나올 수 있는 후보들까지 포함하여 ground truth를 확장하는 2가지 방법을 제안하고 테스트를 수행하였다. 첫 번째 자음/모음 confusion matrix를 이용한 방법으로는 오인식 가능 후보와 OCR 오인식 결과가 겹치는 경우는 발생하지 않았지만, 두 번째 POI 이름을 이용한 방법으로는 2.7배 정도의 성능향상을 확인할 수 있었다.

대해서 confusion matrix를 구해낸다. 이 confusion matrix에서 특정 확률 이상으로 빈번하게 오인식 되는 자음과 모음들을 조합하여 POI 이름 오인식 결과로 나올 수 있는 후보들로 생성해내고 이를 original DB에 추가하여 최종적으로 expanded DB를 만들어낸다.

#### 1. 서 론

이동로봇과 사람 같은 이동이 가능한 물체가 이동시에 현재의 위치를 파악하기 위해서는 주로 GPS 정보가 이용된다. 하지만 실제로 주로 자율주행, 길안내 등의 서비스가 이루어지는 도심, 또는 혼잡 지역에서는 노이즈 및 음영현상 등으로 인하여 GPS 정보가 부정확하거나 이용할 수 없는 상황이 빈번하게 발생한다. 따라서 이동체의 위치를 추정하기 위해서는 단순히 GPS 정보에만 의존하기보다 상호보완적으로 사용할 수 있는 추가적인 정보들을 이용하는 것이 필요하다. 이러한 추가적인 정보들 중의 하나로 POI(Point Of Interest) 정보를 이용할 수 있다. 이동체의 현재 위치에서 획득된 영상 내에 POI의 간판이나 표지판들이 존재하고 여기에 POI의 이름(e.g. “스타벅스” 등)이 적혀 있다면, OCR(Optical Character Recognition) 수행결과로 문자열을 도출하고 이를 이미 알고 있는 POI들의 이름 정보와 매칭시킬 수 있다. 그리하여 매칭된 POI의 위치정보를 이용하여 이동체의 현재 위치를 역으로 추정할 수 있을 것이다. 그런데 OCR 수행결과로 도출된 문자열을 가지고 그것에 해당하는 POI의 이름을 매칭하는 작업은 쉽지 않은 문제이다. 왜냐하면 OCR 수행결과에는 오타자(e.g. “스타벅스” 등) 등의 오류가 빈번하게 발생하기 때문이다. 이를 근본적으로 해결할 수 있는 방법은 OCR 정확도를 높여서 성능 자체를 향상시키는 방법이겠지만, 이는 현실적으로 한계가 있고 현장에서는 주로 OCR 수행결과에 NLP(Natural Language Processing) 등의 후처리를 하여 사용하는 방법을 이용한다. 결과론적으로는 이 방법도 OCR의 성능을 높이는 방식의 하나라고 볼 수 있다. 하지만 문제 해결에 대한 관점을 바꾼다면 새로운 방식을 생각해 볼 수도 있을 것이다.

본 논문에서는 OCR의 성능을 향상시켜서 문제를 해결하지 않고, OCR 오인식 결과로 나올 수 있는 문자열들까지 POI 이름으로 가지고 있어서 문제의 정답지를 확장(expanded DB) 시키는 방식으로 문제를 해결하는 방법을 제안하고 이를 실제로 테스트하여 결과를 확인하였다. .

#### 2. 본 론

##### 2.1 Expanded DB 생성

본 논문에서는 원래 기존에 알고 있는 POI들의 이름을 original DB라고 하고, original DB에 OCR 수행결과로 오인식 되어 나올 수 있는 후보들까지 포함하여 ground truth를 확장한 DB를 expanded DB라고 명명한다. expanded DB를 만드는 방법은 크게 2가지로 나뉜다. 첫 번째는 한글의 자음과 모음들 각각에 대한 confusion matrix를 만들고 확률값이 높은 자음과 모음들을 조합하여 POI 이름에 대해 오인식 결과로 나올 수 있는 후보들로 생성해내는 방법이고, 두 번째는 original DB의 POI 이름들이 OCR 수행결과로 오인식 될 수 있는 경우들을 조사하고 이것들을 original DB에 추가하여 expanded DB를 만드는 방법이다.

##### 2.1.1 자음/모음 confusion matrix 이용

한글의 자음/모음 confusion matrix를 이용하여 expanded DB를 생성하는 방법을 구체적으로 살펴보면 다음과 같다. 먼저 자음(ㄱ, ㄴ, ㄷ, ..., ㅎ)과 모음(ㅏ, ㅑ, ㅓ, ..., ㅣ)을 각각 original image로 생성해내고 이것들에 대해서 여러 가지 변형(variant)을 줘서 variant image들을 만들어낸다. 변형에는 다양한 font, 시점 왜곡(perspective distortion), 영상 일부 잘림(cropping), 영상 일부 가려짐(occlusion), 영상 흐림(blurring), 글자 배경 텍스처(색상) 변경 등이 있을 수 있다. 그리고 나서 이렇게 생성된 original & variant image들을 각각의 자음과 모음으로 OCR이 어떻게 인식하였는지에

〈그림 1〉 자음/모음 confusion matrix 예시

##### 2.1.2 POI 이름 이용

original DB가 이미 가지고 있는 POI 이름들을 이용하여 expanded DB를 생성하는 방법을 구체적으로 살펴보면 다음과 같다. 먼저 original DB에 저장되어 있는 POI 이름들을 가지고 original image를 생성해내고 이것들에 대해서 여러 가지 변형을 줘서 variant image들을 만들어낸다. 변형에는 앞선 방법에서처럼 다양한 font, 시점 왜곡, 영상 일부 잘림, 영상 일부 가려짐, 영상 흐림, 글자 배경 텍스처(색상) 변경 등이 있을 수 있다. 그리고 나서 이렇게 생성된 original & variant image들에 OCR을 수행하였을 때에 오인식되어 나오는 문자열 결과별로 오인식하는 확률들을 구해낸다. 이 확률들 중에서 특정값 이상으로 빈번한 경우의 문자열들을 POI 이름 오인식 결과로 나올 수 있는 후보들로 생성해내고 이를 original DB에 추가하여 최종적으로 expanded DB를 만들어낸다.

Ground Truth OCR	스타벅스	Ground Truth OCR	아가방	Ground Truth OCR	스타벅스	Ground Truth OCR	네거리
스타 벅스	0.81	아 가 방	0.83	스타 벅스	0.74	네 거 리	0.85
스타 벅스	0.15	아 기 방	0.15	스타 벅스	0.13	내 거 리	0.11
스타 벅스	0.02	아 가 빙	0.01	스타 벅스	0.08	비 거 리	0.03
스타 벅스	0.02	아 가 뺑	0.01	스타 벅스	0.05	네 괴 리	0.01
...	...	...	...	...	...	...	...

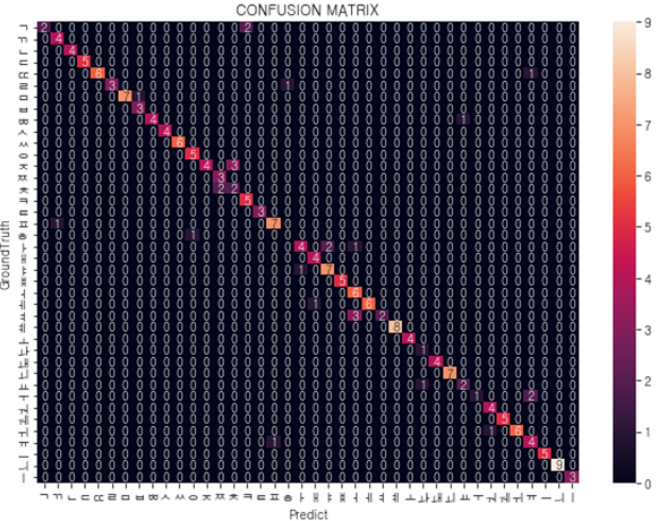
〈그림 2〉 POI 이름 오인식 OCR 결과 예시

##### 2.2 실험 및 결과

##### 2.2.1 자음/모음 confusion matrix 이용

자음/모음 confusion matrix를 이용한 방법에 대해서는, 총 800개의 자음/모음 image(original image: 40개 / variant image: 760개)들을 생성[1]하여 파이썬에서 Keras 라이브러리[2]를 이용하여 confusion matrix를 구하고 2 이상의 빈도로 오인식 되는 자음과 모음들을 조합하여 오인식 가능 후보로 만들어서 실험을 진행하였다. 하지만 예상과는 다르게 오인식 가능 후보와

OCR 오인식 결과가 겹치는 경우는 발생하지 않았다. 이는 자음/모음 각각의 오인식 보다 조합된 상태의 오인식이 중요한 요소이기 때문에 발생한 현상이었던 것으로 판단된다.



〈그림 3〉 자음/모음 confusion matrix

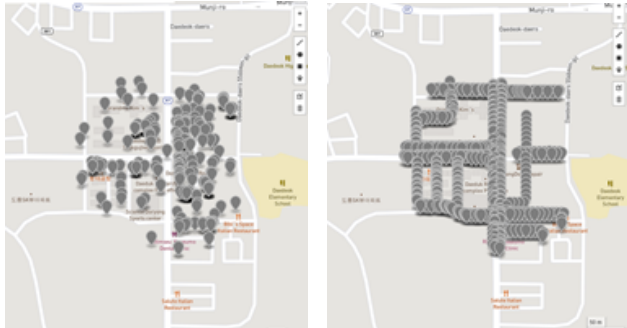


〈그림 4〉 테스트 image

coordinates	predicted_labels
[839.0, 45.0, 1414.0, 185.0]	커피티처
[189.0, 62.0, 709.0, 199.0]	바디샵
[18.0, 73.0, 154.0, 188.0]	더
[494.0, 273.0, 644.0, 420.0]	바
[89.0, 280.0, 420.0, 420.0]	고디
[742.0, 303.0, 1458.0, 421.0]	세제리니스페찌
[34.0, 531.0, 696.0, 673.0]	할리스커피
[765.0, 533.0, 1032.0, 634.0]	네이처
[1077.0, 537.0, 1456.0, 643.0]	리퍼블릭
[1232.0, 777.0, 1451.0, 897.0]	커피
[240.0, 780.0, 668.0, 913.0]	올리브영
[759.0, 782.0, 1198.0, 906.0]	스타벅스

〈그림 5〉 테스트 결과

### 2.2.2 POI 이름 이용



〈그림 6〉 테스트 지역의 POI와 street-view 분포

POI 이름을 이용한 방법에 대해서는, 특정 지역(lat:36.385651, lon:127.378635, radius:200.0)을 선정하여 해당 지역 내의 POI 이름들을 이용하여 image(original image: 267개 / variant image: 2403개)들을 생성하고 이를 오인식한 결과들을 오인식 가능 후보로 만들어서 실험을 진행하였다. 테스트 지역 내의 587개의 street-view image들을 OCR을 수행한 결과, original DB에서는 3개의 POI만 매칭되었지만 expanded DB에서는 8개가 매칭됨을 확인할 수 있었다.

### 3. 결 론

원래 기존에 알고 있는 POI들의 이름에 OCR 수행결과로 오인식 되어 나올 수 있는 후보들까지 포함하여 ground truth를 확장하는 2가지 방법을 제안하고 테스트를 수행하였다. 첫 번째 자음/모음 confusion matrix를 이용한 방법으로는 예상과 다르게 오인식 가능 후보와 OCR 오인식 결과가 겹치는 경우는 발생하지 않았지만, 두 번째 POI 이름을 이용한 방법으로는 2.7배 정도의 성능향상을 확인할 수 있었다.

### Acknowledge

이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No. 2019-0-01309, 불확실한 지도 기반 실내·외 환경에서 최종 목적지까지 이동로봇을 가이드 할 수 있는 AI 기술 개발)

### [참 고 문 헌]

[1] 임재호, 최성록, 이재영, “한글 OCR 데이터셋 생성 기법에 대한 연구”, CICS2020, pp. 67-68, 2020  
 [2] <https://keras.io/>