

2018

하계종합학술대회

2018 IEIE SUMMER CONFERENCE

**Artificial Intelligence led by IEIE:
Now and the Future**

2018년 6월 27일(수) ~ 29일(금) | 롯데호텔 제주(중문)

주최 : 사단법인 대한전자공학회

후원 : 한국과학기술단체총연합회, 해동과학문화재단, SK하이닉스,
삼성전자, LG전자, 전자부품연구원, 한국전자통신연구원,

 제주특별자치도,  |사| 제주컨벤션뷰로

협찬 : 현대자동차, 현대모비스, 현대오트론, 전자정보연구정보센터,
케이던스코리아, 이니프로, 서울대학교 반도체공동연구소



대한전자공학회
The Institute of Electronics and Information Engineers

361. 딥러닝에서 Sparse Autoencoder를 이용한 데이터 feature representation 방법과 문제점 및 개선방안 CFP-486	
김현수, 서승우(서울대학교)	1123

컴퓨터 소사이어티 (포스터)

362. 휴대폰 분실방지용 송수신기 펌웨어 및 HW 모듈 개발 CFP-002	
황기현, 김용우(동서대학교)	1126
363. 효율적인 건설현장 관리를 위한 작업 공정 모니터링 시스템 설계 및 구현 CFP-009	
이좌형, 박정호, 장병태(한국전자통신연구원)	1129
364. 영상 기계학습을 위한 Ground Truth 메타데이터 구조 CFP-015	
추유식, 양창모(전자부품연구원)	1133
365. 오픈소스 스크립트 언어 엔젤스크립트를 이용한 개발/유지보수간 효율적인 시험엔진 구현 CFP-018	
김지성(한화시스템)	1137
366. HDCA 시스템을 위한 고속 PCB 설계 CFP-021	
권원옥, 김영우(한국전자통신연구원)	11470
367. 섬유기반 센서정보를 활용한 실시간 피트니스 자세교정 서비스를 위한 i-Fitness 플랫폼 구현 CFP-044	
오예지, 송재종(전자부품연구원)	1143
368. 함정 컴퓨팅체계를 위한 시스템 자원 관리 방안 연구 CFP-050	
이종현, 김현실, 조태현, 이인수(경북대학교)	1147
369. 조명 변화에 강인한 농작물 병해충 영상 분류 알고리즘 CFP-055	
김미선, 우현준, 한동일(세종대학교)	1149
370. 시간 효율성을 고려한 최적화 신경망 구조 탐색 CFP-056	
김미선, 문재영, 한동일(세종대학교)	1153
371. IoT 기반 스마트캐비닛용 게이트웨이 개발 CFP-067	
권오훈, 정민우, 이상훈, 이주희(경북IT융합산업기술원)	1157
372. 딥 러닝을 이용한 명함 인식 시스템 CFP-069	
석정환, 윤준서, 박창우, 김동호(동국대학교)	1160
373. 딥러닝을 이용한 박물관 전시품 인식 어플리케이션 개발 CFP-070	
조은지, 임정수(한국교통대학교)	1164
374. 비주얼 서보잉 물체 조립을 위한 대상 물체 외곽 성분 검출 방법 CFP-075	
조재민, 김계경(한국과학기술연합대학원대학교)	1168

포스터

컴퓨터 소사이어티



딥 러닝을 이용한 명함 인식 시스템

*석정한, *윤준서, **박창우, ***김동호

*동국대학교 컴퓨터공학과

**동국대학교 경영정보학과

***동국대학교 융합소프트웨어교육원

e-mail : rudebono@gmail.com, yun6686@gmail.com, ckddn3310@naver.com, dongho.kim@dgu.edu

A system for business card recognition with deep learning

*Jeonghan Seok, *Junseo Yun, **Changwoo Park, ***Dongho Kim

*Department of Computer Science and Engineering

**Department of Management Information System

***Convergence Software Institute

Dongguk University

Abstract

본 연구에서는 기존의 명함 인식 시스템의 한계점을 소개하고 이를 보완할 방법으로 딥 러닝을 이용한 명함 인식 시스템을 제시 및 구현하여 실험한다. 명함 인식 시스템의 구조는 크게 이미지 속 문자를 탐지 및 인식하는 전반부와 인식된 문자를 명함에서 사용된 개체별로 인식하는 후반부로 나누어진다. 본 논문에서는 명함 인식 시스템의 전반부와 후반부와 관련 있는 인공지능 모델인 CTPN, CRNN 그리고 BLSTM CNNs CRF를 소개하고 이를 이용하여 명함 인식 시스템을 제시한다. 실제 구현에서는 명함 인식 시스템의 전반부에 해당하는 문자 탐지 및 인식 기능은 구글사의 이미지 처리 서비스인 구글 클라우드 비전을 이용하며, 후반부에 해당하는 개체명 인식 모델은 오픈 소스 라이브러리인 텐서플로를 이용하여 BLSTM CNNs CRF 모델을 구현한다. 개체명 인식 모델은 조직명, 성명, 주소 등 명함에서 사용하는 개체별 데이터를 조합하여 생성한 10,000,000개의 데이터를 학습한다. 이렇게 구현한 명함 인식 시스템을 실제 명함 이미지 1,000개에 대해 실험한 결과 후반부에 해당하는 개체명 인식의 인식률이 낮았으며, 이 문제점에 대해 향후 연구 방향을 제시한다.

I. 서론

명함은 사회에서 일상적으로 사용하는 통신 및 저장 매체로서 그 역할이 매우 중요하다. 하나의 작은 명함에서 조직명, 성명, 주소, 전화 번호, 직책, 전자 메일 등과 같은 중요한 개인 정보를 찾을 수 있다. 하지만 수많은 명함에서 자신이 얻고자 하는 정보를 찾기란 쉽지 않다. 이러한 정보를 데이터베이스에 수동으로 입력하여 명함을 관리 할 수 있지만 불편하고 시간이 오래 걸린다. 따라서 명함의 정보를 자동으로 추출하는 시스템을 개발하는 것이 유용하며 이에 대해 이미 많은 연구가 진행되었다. 기존의 명함 인식 시스템은 특정한 알고리즘을 바탕으로 한다. 하지만 이러한 접근 방법은 다양한 형태의 명함을 모두 처리할 수 없는 한계점을 지닌다.

본 연구에서는 명함 인식 시스템에 대한 기존의 관련 연구의 접근 방법과 한계점을 소개하고, 딥 러닝을 이용한 새로운 접근 방법을 이용하여 명함 인식 시스템을 제시한다. 또한, 이를 구현하여 실제 명함에 대해 실험 후 결과를 도출하고 이를 분석하여 향후 연구 계획을 제시한다.

II. 본론

2.1 관련 연구

명함 인식 시스템에 대한 기존의 연구는 모바일 환경에서의 명함인식 성능 향상에 관한 연구[1]에서 볼 수 있듯이 특정한 알고리즘을 이용하여 이미지 처리를 통해 이미지 속 문자를 추출하고 추출된 문자는 형태소, 품사, 구절, 구문 등 다양한 문자 분석을 통해 개체명을 인식하여 명함에서 사용된 필드별로 분류한다. 이러한 특정 알고리즘과 통계를 바탕으로 하는 접근 방법은 특정한 데이터에 대해서는 좋은 성능을 낼 수 있지만, 사람이 직접 데이터를 분류하고 특징을 추출하여 알고리즘을 도출하는 어려움과 이미지와 문자의 다양성을 모두 고려할 수 없는 한계점을 갖고 있다. 모든 경우의 수를 고려하기 쉽지 않기 때문이다. 따라서 본 연구에서는 이러한 한계점을 해결하고자 인공지능 기술을 이용해 축적된 데이터를 스스로 분석한 후 답을 내는 딥 러닝 방식으로 접근을 시도한다.

2.2 딥 러닝을 이용한 명함 인식 시스템

딥 러닝 분야에서 이미지 속 문자를 추출하는 방법으로 문자 탐지와 인식에 관한 많은 연구가 있으며, 대표적으로 CTPN(Connectionist Text Proposal Network)[2]을 이용한 문자 탐지와 CRNN(Convolutional Recurrent Neural Networks)[3]을 이용한 문자 인식이 있다. 또한, 문자 속의 개체명을 인식하는 방법으로 많은 연구가 있으며, 대표적으로 BLSTM CNNs CRF(Bidirectional Long Short-Term Memory Convolutional Neural Networks Conditional Random Field)[4]을 이용한 개체명 인식이 있다.

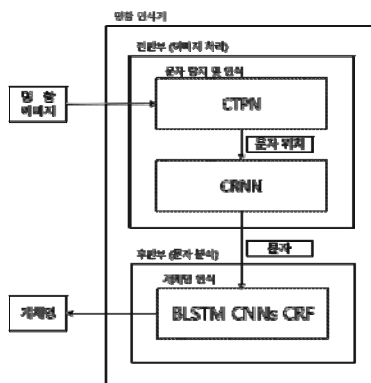


그림 1. 명함 인식 시스템 구조

명함 인식 시스템의 구조는 그림 1에서 볼 수 있듯이 전반부와 후반부로 크게 두 가지로 구성되어 있다. 전반부는 이미지를 처리하여 이미지 속 문자를 인식하

는 영역이다. 이미지 속 문자의 위치를 탐지하는 CTPN과 탐지된 영역에서 문자를 인식하는 CRNN 모델을 사용한다. 후반부는 추출된 문자를 명함에서 사용된 개체별로 인식하는 BLSTM CNNs CRF 모델을 사용한다.

명함 속에는 문자 이외 조직 로고나 사람 얼굴 등과 같은 이미지가 포함될 수 있다. 문자 인식에 앞서 명함 속 문자의 위치를 탐지하는 것이 필요하다. CTPN은 전체 이미지 중 문자 영역을 기반으로 학습한다. 이미지 속 문자 영역을 탐지하면 해당 영역의 문자를 인식하는 과정이 필요하다. CRNN은 CTPN과 다르게 전체 이미지가 아닌 문자 영역의 이미지를 기반으로 학습한다. 이미지 속 문자를 추출하면 이를 명함 속에 존재하는 필드별로 분류하는 개체명 인식 과정이 필요하다. 최근 개체명 인식 연구에서 우수한 성능을 보여주고 있는 모델은 BLSTM CNNs CRF 모델이며 문자와 해당 개체명을 기반으로 학습한다. 앞서 소개한 모델들의 특징은 CNN을 이용하여 이미지나 문자의 특징을 추출하고 BLSTM을 이용하여 특징을 연결시킨다는 점이다.

III. 구현

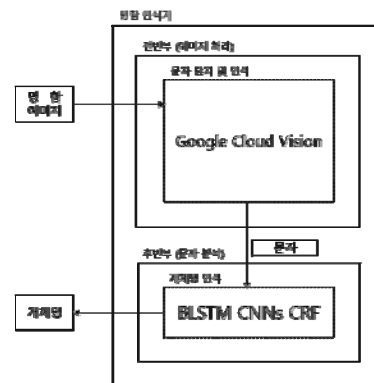


그림 2. 명함 인식 시스템 구현

본 연구에서 제시한 명함 인식 시스템의 구조 중 전반부에 해당하는 이미지 처리 모델은 학습에 필요한 시간과 자원이 부족하여 그림 2와 같이 구글 클라우드 비전(Google Cloud Vision)[5]으로 대체 하였다.

구글 클라우드 비전은 구글사에서 제공하는 이미지 처리 API(Application Programming Interface)이다. 구글 클라우드 비전은 이미지 안에서 객체와 얼굴을 감지하고, 이미지 카탈로그에서 메타데이터를 제작하거나, 불쾌감을 주는 콘텐츠를 검토하는 등 다양한 이미지 처리를 제공한다. 본 연구에서는 이러한 다양한 기능 중

에서 명함 인식 시스템 구조의 전반부에 해당하는 이미지 속 문자를 탐지하고 인식하는 기능을 사용한다.

BLSTM CNNs CRF는 정상근 박사의 모두를 위한 인공지능[6]에서 공개된 텐서플로우(Tensorflow)[7]로 구현한 모델을 명함 개체명 인식 모델에 맞게 수정하여 구현하였다. 그리고 한국콘텐츠미디어의 한국 SMTp 2016[8]에서 회사명 1,000,000개, 인명 180,000개, 주소 1,300,000개, 직책 및 직급 509개, 홈페이지 90,000개를 추출하고, 휴대폰 번호, 전화 번호, 팩스 번호를 대한민국 전화 번호 체계에 맞게 생성하여 10,000,000개의 학습 데이터를 생성하고 모델을 학습하였다.

IV. 실험

명함 인식 시스템은 크게 이미지 처리인 전반부와 문자 분석인 후반부로 나눌 수 있으며, 본 연구에서는 구글 클라우드 비전과 BLSTM CNNs CRF 모델을 이용하여 구현하였다. 구현한 명함 인식 시스템은 이미지에서 인식한 문자를 기반으로 문자 분석을 하므로 전반부의 성능은 후반부에 영향을 미칠 수 있다. 그래서 본 실험에서는 명함 1,000장에 대해 구글 클라우드 비전의 문자 탐지 및 인식에 대한 결과로 BLSTM CNNs CRF의 개체명 인식률을 분석하였다.

4.1 문자 탐지 및 인식

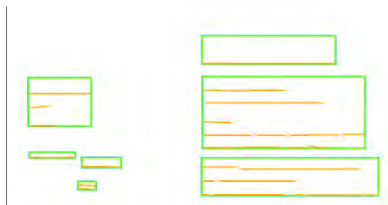


그림 3. 영역 탐지 결과



그림 4. 문자 탐지 결과

구글 클라우드 비전에서 이미지 속 문자 영역을 탐지할 때 그림 3과 같이 먼저 하나의 이미지를 여러 영역을 나눈다. 문자가 서로 다른 높이에 존재하기 때문이다.

unIPoint 송훈섭 이사 연구소 (주)유니포인트
http://www.unipoint.co.kr 07024 서울특별시 동작구 남
부순환로 2049 (사당동, 네오티스 빌딩 2층) Microsoft
Partner Ahntab E-mail: hssong@unipoint.co.kr M.
010-3121-5855 T. 02-6676-5516 F. 02-6676-5599
Buaniess Fanne

표 1. 문자 인식 결과

unIPoint
송훈섭
이사
연구소
(주)유니포인트
http://www.unipoint.co.kr
...

표 2. 구분된 영역 문자 인식 결과

영역을 분리 후 그림 4와 같이 영역 속에 존재하는 문자를 탐지한다. 구글 클라우드 비전에서는 이렇게 탐지된 문자를 인식하여 표 1과 같이 줄 바꿈이 없는 하나의 문장이나 표 2와 같이 그림 3에서 나누었던 영역별로 나뉜 문장으로 인식 결과를 제공한다.

항 목	인식률
문자 탐지	97%
문자 인식	95%

표 3. 구글 클라우드 비전 성능

명함 1,000장을 이용하여 구글 클라우드 비전의 다양한 기능 중 문자 탐지와 인식에 대한 성능 평가를 하였다. 문자 탐지 기능에서 작은 글씨나 특정 회사의 로고를 인식하지 못하는 한계점이 있으며 97%의 문자 탐지 성능을 보였고, 문자 인식 기능에서는 회사의 로고, 작은 글씨, 띄어쓰기 그리고 특수문자가 올바르게 인식되지 않는 한계점이 있으며 95%의 문자 인식 인식률을 보였다.

4.2 개체명 인식

문 자	개체명	확 른
송	B-NM	0.999997
훈	I-NM	0.999999
섭	I-NM	0.999998
	O	0.999997
이	B-PS	0.818332
사	I-PS	0.986787
	O	0.805898
(B-OG	0.912391
주	I-OG	0.540363
)	I-OG	0.514145
유	I-OG	0.568878
니	I-OG	0.504413

표 4. 개체명 인식 결과

명함 이미지 속에서 인식된 문자를 분류하는 과정에 대해 BLSTM CNNs CRF 모델의 성능을 평가한다.

분류 항목은 조직명, 직책 및 직급, 성명, 주소, 휴대폰 번호, 전화 번호, 팩스 번호, 홈페이지 주소로 총 8개이다. BLSTM CNNs CRF 모델에 추출한 문자를 입력하면 표 4에서 보는 것처럼 각 문자에 해당하는 분류 항목과 추측 확률을 결과로 받는다.

항 목	인식률
조직명	48%
직책 및 직급	22%
성명	42%
주소	62%
휴대폰 번호	42%
전화 번호	43%
팩스 번호	48%
홈페이지 주소	55%

표 5. 한 문장에 대한 BLSTM CNNs CRF 성능

항 목	인식률
조직명	63%
직책 및 직급	31%
성명	58%
주소	72%
휴대폰 번호	68%
전화 번호	63%
팩스 번호	62%
홈페이지 주소	82%

표 6. 구분된 문장에 대한 BLSTM CNNs CRF 성능

명함 1,000장을 구글 클라우드 비전을 통해 이미지 속 문자를 인식하고, 인식된 문자에 대해 BLSTM CNNs CRF 모델의 개체명 인식 성능을 평가 하였다. 개체명 인식 과정에서 각 개체별로 중복된 분류 항목이 나올 수 있으며 실험에서는 중복된 항목 중 추측 확률이 제일 높은 항목을 선택하여 측정하였다. 표 5는 줄 바꿈이 없는 하나의 문장에 대해, 표 6은 줄 바꿈이 있는 구분된 문장에 대한 개체명 인식에 결과이다. 하나의 문장 대신 구분된 문장을 이용하면 약 10%~30%의 인식률 향상을 보였다.

V. 결론 및 향후 연구 방향

본 연구에서는 딥 러닝을 이용한 명함 인식 시스템을 제시 및 구현하고 실험을 통해 성능을 분석하였다. 이미지 속 문자를 인식하는 과정에서 구글 클라우드 비전을 사용하였으며 작은 글씨나 특정 회사의 로고를 인식하지 못하는 한계점을 가지고 있으나 95% 이상의 높은 인식률을 보였다. 개체명 인식 과정에서 BLSTM CNNs CRF 모델을 사용하였으며 줄 바꿈이 없는 하

나의 문장보다 구분된 문장에서 약 10%~30%의 인식률 향상을 보였다. 하지만 두 실험 모두 90% 이상의 인식률을 나타내지 못하였는데 이는 학습의 부족이나 실제 명함의 데이터를 기반으로 학습시키지 않고 사용자에 의해 생성된 데이터를 기반으로 학습을 시킨 이유로 추측하고 있다. 더욱 많은 데이터를 이용하거나 실제 명함 데이터를 이용하여 개체명 인식 모델을 학습하여 본 연구에서 제시한 명함 인식 시스템의 인식률과 비교하는 것을 향후 연구 방향으로 제시한다.

감사의 글

본 논문은 과학기술정보통신부 및 정보통신기술진흥센터의 재원으로 SW중심대학지원사업의 지원을 받아 수행된 연구임. (No. 2016-0-00017)

참고문헌

- [1] 김차중, 신현섭, “모바일 환경에서의 명함인식 성능 향상에 관한 연구”, 한국정보통신학회논문지 (J.. Korea Inst. Inf. Commun. Eng.) February, 2014, Vol. 18, No. 2, pp 318-328.
- [2] Zhi Tian, Weilin Huang, Tong He, Pan He and Yu Qiao, “Detecting Text in Natural Image with Connectionist Text Proposal Network”, European Conference on Computer Vision, 2016, pp 56-72.
- [3] Baoguang Shi, Xiang Bai and Cong Yao, “An End-to-End Trainable Neural Network for Image-based Sequence Recognition and Its Application to Scene Text Recognition”, IEEE Transactions on Pattern Analysis and Machine Intelligence, November 2017, pp 2298-2304.
- [4] Xuezhe Ma and Eduard Hovy, “End-to-end Sequence Labeling via Bi-directional LSTM-CNNs-CRF”, ACL 2016.
- [5] <https://cloud.google.com/vision/>
- [6] <http://www.hugman.re.kr/>
- [7] <https://www.tensorflow.org/>
- [8] <http://www.emnb.co.kr/>