

# 도면 문자 인식 성능 향상을 위한 전처리와 결합한 변형된 VGG 심층학습 모델

신희란\*, 박장식\*, 송종관\*  
\*경성대학교 전기전자공학과  
e-mail : heeran3646@gmail.com

## Modified VGG Deep Learning Model with Pre-processing for Improving Optical Character Recognition of Machine Drawings

Hee-ran Shin\*, Jongkwan Song\*, Jangsik Park\*  
\*Dept of Electronic Engineering, Kyungsung University

### 요 약

본 논문에서는 전처리와 결합한 변형 VGG 모델을 이용하여 기계도면에서의 도형과 지시선에 의한 오인식을 해결하고 문자 인식 성능을 향상시킨다. 제안하는 도면 문자 인식은 도형과 지시선을 제거하기 위하여 수학적 형태학 전처리를 수행하고, 작은 숫자를 인식하기 위하여 변형된 VGG 심층학습 모델을 적용한다. 변형된 VGG 심층학습 모델은 기존의 VGG-16 모델의 계층을 7계층 줄인 것이다. 제안한 도면 문자 인식 방법에 대한 성능 평가 결과 기계 부품 서비스 도면 100장에 대하여 인식률 95.57%, 정확도 92.82%의 결과를 얻었다.

### 1. 서론

자동화 시스템을 갖춘 생산 체계에서는 생산 및 공정, 빌딩의 경우 각종 설비들을 원격으로 제어 및 감시, 사무 및 회계 또한 인간이 수기로 작성하던 데이터의 일괄처리가 있다. 광학 문자 인식(OCR, Optical Character Recognition)은 사람의 필기체 및 인쇄된 문자를 시스템이 읽을 수 있는 문자 코드로 변환하는 기술이다. 이미 광학 문자 인식 기술은 주차장에서의 차량 번호판 인식, 처방전 변환 등 많은 곳에서 사용되고 있는 만큼 관련 연구들 또한 활발하다[1]. 기계 산업 분야 또한 효율적인 부품 공급을 위해 기존 부품도면의 문자 인식이 필요하다. 기업에서는 제품 제작을 위한 설계도면, 부품 정보 유출을 막기 위한 배포용으로 제공하는 서비스 도면 두 가지를 제작한다. 서비스 도면의 부품과 명칭을 연결할 때 인력이 필요하며 많은 시간과 비용이 소요된다. 구글(Google)에서 개발한 Tesseract OCR은 워드 문서 및 글자를 변환하는데 좋은 인식률을 나타낸다[2]. 그러나, 도면의 경우, 선과 도형이 문자로 오인식 되는 문제가 발생한다..

본 논문에서는 서비스 기계 부품에서의 인식률을 향상하기 위해 전처리와 결합된 심층학습 기반 문자 인식 방법을 제안한다. 수학적 형태학(Mathematical morphology) 전처리 과정을 거쳐 기계 부품의 지시선 및 도형, 즉 오인식 가능성이 있는 요소를 제거하고, 심층학습을 통하여 추론한다.

### 2. 제안하는 OCR

서비스 부품 도면에서 일반적인 문서와 달리 부품 도형과 지시선이 숫자로 오검출 되기 때문에 OCR 정확도가 떨어진다. 서비스 부품 도면에 인식해야 할 숫자와 도형 및 지시선의 두께 차이를 이용하여 수학적 형태학 필터링을 하여 제거한다. 제거된 이미지 영상에서 학습된 심층학습 모델을 이용하여 추론하고 각 서비스 부품 도면마다 인식률과 정확도를 측정한다. 서비스 부품 도면에서 인식해야 할 숫자를 측정하여(Ground Truth) 텍스트 파일로 제작하여 학습 데이터셋을 구성한다.

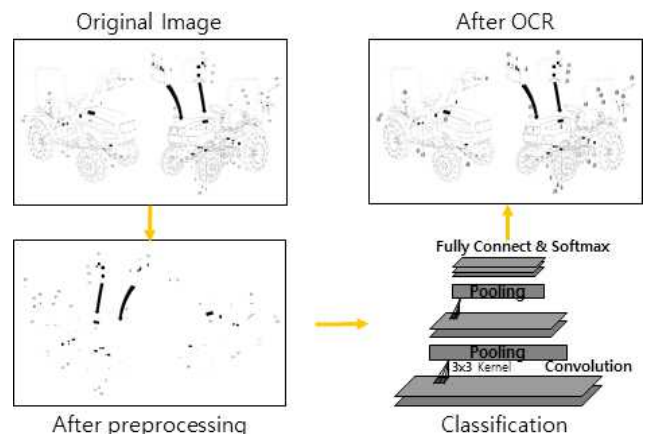


그림 1. 제안하는 심층학습 OCR 구조

#### 2.1 수학적 형태학 필터 전처리

수학적 형태학 필터 연산은 특정 마스크를 이용해 이미지 객체를 축소 및 확장해 객체 변형을 유도한다. 기본적으로 침식(erosion)과 팽창(dilation)이 있으며, 두 연산의 복합 구성으로 침식 이후 팽창을 수행하는 열림(Opening)과 팽창필터 침식필터 순으로 수행하는 닫힘(Closure)이 있다[3-4].

본 논문에서는 열림 연산을 사용했다. 처음으로 침식 연산을 사용해 서비스 부품 도면의 도형과 선을 제거했다. 인식해야 할 문자는 부품 도형 및 선의 굵기보다 상대적으로 두껍다. 이 점을 이용하여 3x3 크기의 십자가 모양 마스크를 사용해 침식 연산을 수행하여 최대한 제거한다. 침식 연산으로 인식해야 하는 문자가 손상된 것을 보완하기 위해 팽창 연산을 수행한다.

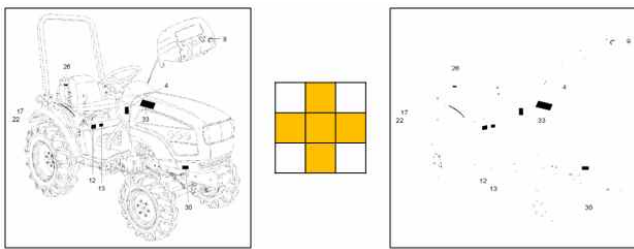


그림 2. 왼쪽부터 원본영상, 사용 마스크, 전처리 결과

## 2.2 심층학습 모델

사용된 심층학습 모델은 7계층(layer)로 구성되어 있으며 VGG-16과 유사한 구조를 가진다. VGG-16은 3x3 크기의 작은 커널을 사용하는 16계층 심층학습 모델이다. 서비스 부품 도면의 문자 이미지 크기는 가로 30~40, 세로 25~30이다. 이 점을 고려하여 7계층 모델은 입력 데이터 크기가 28x28로 작다.[5]

표 2. 사용된 심층학습 모델 구조

layer(type)		size
conv1_1	Convolution	28x28x1
conv1_2	Convolution	28x28x1
pool_1	Avg Pooling	14x14x1
conv2_1	Convolution	14x14x2
conv2_2	Convolution	14x14x2
pool_2	Avg Pooling	7x7x2
fc1	FC	512
fc2	FC	256
fc3	FC	10

## 3. 실험 결과

실험 데이터 베이스는 도면 100장에 대한 인식률과 정확도를 측정하며, 심층학습 모델을 위한 데이터셋은 도면에서 추출한 0부터 9까지 각각 약 3,000개의 이미지를 사용했다. 학습 환경은 우분투(Linux Ubuntu) 16.04 NVIDIA TITAN X 12GB이다. 학습된 모델을 비교 분석하기 위해 .NET C#기반 애플리케이션을 제작했다. 제작

된 애플리케이션은 선택한 폴더 내의 모든 도면 이미지를 스캔하여 인식한 문자 정보의 좌표나 인식된 숫자를 텍스트로 변환하고 창에 시각화 한다. 표1은 도면 100장에 대한 OCR 인식률 및 정확도를 측정한 것이다.

표 1. 서비스 부품 도면 OCR 성능 비교

Models	preprocessing and VGG	only 7 layers model	proposed OCR
Recall	72.50%	96.04%	95.57%
Precision	69.14%	87.01%	92.82%
F1 score	69.31%	91.30%	94.17%

## 4. 결론

본 논문은 기계 산업 분야의 서비스 부품 도면의 문자 인식 개선을 위해 수학적 형태학 연산과 심층학습 모델을 제안했다. 그 결과 서비스 부품 도면 100장에 대한 OCR 결과는 제안된 OCR이 인식률(Recall) 95.57%, 정확도(Precision) 92.82%, F1척도 94.17%를 나타냈다.

## 감사의 글

본 논문은 중소벤처부 S-BOM기반 애프터마켓용 파츠북 서비스 ITEM 시스템 개발과 2018년 BB21+ 사업에 의하여 지원되었음

## 참고문헌

- [1] 김남규, 김동연, 김성우, 권순각 “광학 문자 인식을 통한 단어 정리 방법,” 한국멀티미디어학회논문지, 제 18권, 8호, pp. 943-949, 2015.
- [2] R. Smith “An overview of the Tesseract OCR Engine,” *IEEE Proc. Int. Conf. on Document Analysis and Recognition*, Parana, Brazil, vol. 2, Sept. pp. 629-633, 2007.
- [3] X. C. Yin, X. Yin, K. Huang, and H. W. Hao, “Robust text detection in natural scene images,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 36, pp. 970-983, 2014.
- [4] L. Vincent, “Morphological area openings and closings for greyscale images,” *Proc. NATO Shape in Picture Workshop, Driebergen, The Netherlands*, S pp. 197-208, 1992.
- [5] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *Proc. of Int. Conf. Learning Representations ICLR 2015, San Diego, USA, May. 2015.* <https://arxiv.org/abs/1409.1556>