

Structuring of Pulmonary Function Test Paper Using Deep Learning

Sang-Hyun Jo*, Dae-Hoon Kim*, Yoon Kim*, Sung-Ok Kwon**, Woo-Jin Kim***, Sang-Ah Lee****

*Student, Dept. of Computer Science and Engineering, Kangwon National University, Chuncheon, Korea

*Student, Dept. of Computer Science and Engineering, Kangwon National University, Chuncheon, Korea

*Professor, Dept. of Computer Science and Engineering, Kangwon National University, Chuncheon, Korea

**Professor, Dept. of Medical Bigdata Convergence, Kangwon National University, Chuncheon, Korea

***Professor, Dept. of Internal Medicine and Biomedical Informatics, Kangwon National University, Chuncheon, Korea

****Professor, Dept. of Preventive Medicine, Kangwon National University School of Medicine, Chuncheon, Korea

[Abstract]

In this paper, we propose a method of extracting and recognizing related information for research from images of the unstructured pulmonary function test papers using character detection and recognition techniques. Also, we develop a post-processing method to reduce the character recognition error rate. The proposed structuring method uses a character detection model for the pulmonary function test paper images to detect all characters in the test paper and passes the detected character image through the character recognition model to obtain a string. The obtained string is reviewed for validity using string matching and structuring is completed. We confirm that our proposed structuring system is a more efficient and stable method than the structuring method through manual work of professionals because our system's error rate is within about 1% and the processing speed per pulmonary function test paper is within 2 seconds.

▶ **Key words:** Text Detection, Text Recognition, Image Classification, Data Structuring, Deep Learning

[요 약]

본 논문에서는 문자 검출 및 인식 기술을 활용하여 비정형의 폐 기능 검사지 이미지로부터 연구자를 위한 관련 정보들을 추출하여 정형화하는 방법을 제안한다. 또한 문자 인식 오차율을 줄이기 위한 후처리 방법 또한 개발하고자 한다. 제안하는 정형화 방법은 폐 기능 검사지 이미지에 대해 문자 검출 모델을 사용해 검사지 내에 존재하는 모든 문자를 검출하고, 검출된 문자 이미지를 문자 인식 모델에 통과시켜 문자열을 얻어낸다. 얻어낸 문자열에 대해 문자열 매칭을 이용한 유효성 검토를 진행하고 정형화를 마무리한다. 제안하는 정형화 시스템의 오차율은 약 1% 이내, 검사지 당 처리속도는 2초 이내로 전문인력의 수작업을 통한 정형화 방법보다 더 효율적이고 안정적인 방식이라는 것을 확인할 수 있다.

▶ **주제어:** 글자 검출, 글자 분류, 이미지 분류, 데이터 정형화, 딥러닝

-
- First Author: Sang-Hyun Jo, Corresponding Author: Sang-Ah Lee
 - *Sang-Hyun Jo (josh3255@kangwon.ac.kr), Dept. of Computer Science and Engineering, Kangwon National University
 - *Dae-Hoon Kim (eogns0707@naver.com), Dept. of Computer Science and Engineering, Kangwon National University
 - *Yoon Kim (yooni@kangwon.ac.kr), Dept. of Computer Science and Engineering, Kangwon National University
 - **Sung-Ok Kwon (kamelon@hanmail.net), Dept. of Medical Bigdata Convergence, Kangwon National University
 - ***Woo-Jin Kim (pulmo2@kangwon.ac.kr), Dept. of Internal Medicine and Biomedical Informatics, Kangwon National University
 - ****Sang-Ah Lee (sangahlee@kangwon.ac.kr), Dept. of Preventive Medicine, Kangwon National University School of Medicine
 - Received: 2021. 11. 04, Revised: 2021. 11. 16, Accepted: 2021. 11. 22.

I. Introduction

한국의 연간 폐 기능 검사자의 수는 800,000명에 육박한다. 하지만 대부분의 병원들은 환자의 폐 기능 검사결과를 이미지 형태로 보관하기 때문에 관리 및 활용에 어려움을 겪고 있다. 폐 기능 검사지는 환자의 이름, 생년월일, 키, 몸무게를 비롯한 개인정보부터 검사항목, 수치, 예측값과 검사소견 등과 같은 의료정보를 포함하고 있다. 검사지에는 평균적으로 110개 내외의 단어가 쓰이며 검사장비나 검사항목에 따라 다양한 출력 형태를 가진다. 이러한 검사지의 특징들로 인해 전문화된 인력이 수작업으로 정형화를 진행 할 경우에는 많은 비용과 시간이 소모되며 5~10% 정도의 오차율을 보인다.

본 논문에서는 이런 문제점을 극복하기 위해 딥러닝을 이용한 폐 기능 검사지의 정형화 시스템을 제안한다. 제안하는 정형화 시스템은 네 단계로 진행된다. 첫 번째는 이미지 분류 모델을 이용한 검사지의 분류이다. 본 연구에서는 후처리 과정 중 하나인 문자 교정을 위해서 검사지를 Fig 1과 같이 6가지 종류로 구분한다. 두 번째는 문자 검출 모델을 이용한 검사지의 문자 검출이다. 검출 대상 문자가 휘어지지 않고 반듯한 문자이기 때문에 검출 성능보다는 처리속도에 중점을 두고 모델을 선정하였다. 세 번째는 문자 분류 모델을 이용한 검사지의 문자 분류이다. 마지막으로 후처리 과정을 통해서 변형된 검사항목과 수치를 조정한다. 모델 훈련과 성능평가에는 강원대학교 병원의 2016년도 상반기부터 2020년도 하반기까지의 비식별화된 폐 기능 검사지 39,900장을 이용하였다.

본 논문의 구성은 다음과 같다. 2장에서는 학습에 사용된 데이터셋에 대한 설명과 문자 검출 및 분류 기술에 대해 기술한다. 3장에서는 본 논문에서 제안하는 폐 기능 검사지 데이터의 정형화 방법을 설명한다. 4장에서는 실험결과에 대해서 설명한다. 5장에서는 결론을 도출한다.

II. Preliminaries

1. Dataset

글자 추출을 위한 모델의 학습 및 평가 데이터셋으로는 강원대학교 병원의 폐 기능 검사지 39,900장과 옥스퍼드 대학에서 공개한 합성 텍스트 데이터셋인 MJSynth와 SynthText를 이용하였다.

1.1 MJSynth

MJSynth[1]는 옥스퍼드 대학의 VGG 그룹에서 공개한 합성 텍스트 데이터셋으로 훈련, 검증, 평가를 위한 900만 개의 이미지로 구성되며 9만 개의 영어단어를 포함한다.

1.2 SynthText

SynthText[2]는 80만 개의 이미지를 포함하는 데이터셋으로 800만 개에 가까운 합성 단어로 구성되어있다. 각 합성 단어에 대한 문자열, 경계상자(Bound Box), 단어 및 문자 수준의 정보를 제공한다.

1.3 Pulmonary Function Test Paper

본 연구에 사용되는 폐 기능 검사지 데이터는 강원대학교병원 생명의학연구윤리심의위원회(IRB)의 허가를 획득하였다. (IRB number B-2020-11-003-005).

폐 기능 검사지는 환자의 이름, 성별, 생년월일, 몸무게, 키와 같은 개인정보부터 검사항목, 검사 수치, 예측된 수치 값, 검사소견 등과 같은 의료정보를 포함하고 있다. 폐 기능 검사지의 단어는 검사항목이나 장비에 따라서 적게는 70개부터 많게는 140개 정도가 사용된다. 본 논문의 실험에는 강원대학교 병원의 비식별화된 폐 기능 검사지 39,900장을 사용하였다.

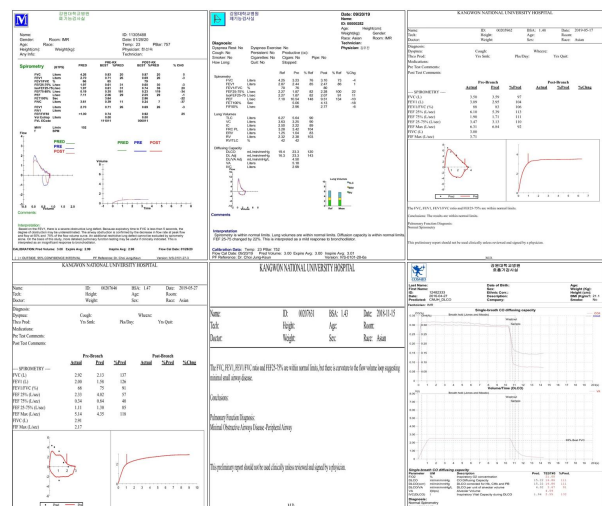


Fig. 1. Type of Pulmonary Function Test Papers

2. Text Detection

문자 검출은 이미지 내에 존재하는 문자를 검출하는 기술이다. 문자 검출 기술은 일반적으로 글자 기반 방식과 단어 기반 방식 두 가지로 구분한다.

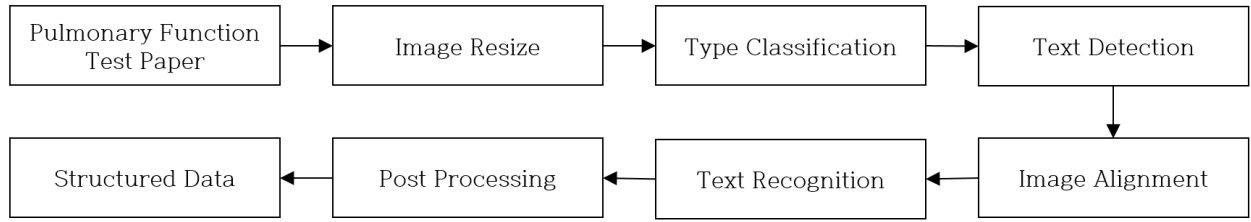


Fig. 2. Process of Data Structuring

2.1 글자 기반 방식

글자 기반 방식은 가장 먼저 글자 후보를 추출하고 필터링, 그룹화를 통해 최종적으로 단어를 검출하는 과정을 가진다. 최근 다양한 데이터셋에 대해서 높은 정확도를 보이는 글자 기반 방식은 CharNet[3], SBD[4], Mask TextSpotter[5] 등의 모델들이 있다. 그러나 글자 기반 방식은 단어 기반 방식에 비해 높은 정확도를 보이지만 전처리와 후처리에 많은 비용이 들어가며 실행시간이 느리다는 단점이 있다.

2.2 단어 기반 방식

단어 기반 방식은 문자 검출을 일반적인 객체 검출과 동일하게 경계상자 회귀문제로 정의한다. DB[6], PMTD[7], PAN[8] 등의 모델이 Word based 구조를 사용한다. 단어 기반 방식은 간단한 구조로 구현과 배포가 용이하지만 글자 기반 방식에 비해서 정확도가 떨어진다는 단점이 있다.

3. Text Recognition

문자 인식은 이미지 속 글자를 인식하는 기술로 지난 몇 년간 딥러닝의 발전과 함께 빠른 속도로 발전했다. 문자 인식 기술은 크게 CTC based methods와 Attention based methods 두 가지 방식으로 구분한다.

3.1 CTC based methods

CTC 기법은 모델의 출력 시퀀스 벡터에 대해서 별도의 조정 없이 정답 레이블을 학습 할 수 있게 해주는 방식이다. 이러한 방식은 구현과 레이블링에 소모되는 비용이 적지만 Attention based methods에 비해 떨어지는 정확도를 보여준다.

3.2 Attention based methods

Attention 기법은 모델이 출력 단어를 예측하는 시점에 이전 시점의 입력이나 디코더 출력을 참고하는 방식이다. Yet Another Text Recognizer는 2D-Attention이 사용된 네트워크 구조로 ICDAR2003, ICDAR2013 그리고 ICDAR2015 데이터셋에 대해서 SOTA를 달성했다[9].

III. The Proposed Scheme

이미지 정형화 프로세스는 입력받은 이미지에 대한 크기 조절, 검사지 타입 분류, 글자 검출, 이미지 정렬, 글자 분류, 후처리 과정 순서로 진행된다.

1. Pre Processing

전처리 알고리즘은 폐 기능 검사지의 분류를 위한 이미지 크기 조절, 그리고 문자 인식 과정에서의 왜곡 및 회전된 문자로 인한 오인식을 방지하기 위한 이미지 정렬 두 가지로 구성된다. 분류 모델의 빠른 처리속도를 위해서 이미지 크기 조절은 224×224 크기로 진행하였다.

이미지의 왜곡 보정을 위해서 아핀 변환을 사용하여 이미지 정렬을 진행한다. 검출된 경계상자 좌표에 회전, 이동 및 크기변환 등을 사용해 사전에 정의한 좌표로 변환하면서 이미지를 정렬시킨다. 임의의 세 좌표인 (x_1, y_1) , (x_2, y_2) , (x_3, y_3) 를 특정 좌표인 (x'_1, y'_1) , (x'_2, y'_2) , (x'_3, y'_3) 로 변환하는 행렬식은 식(1)과 같다.

$$\begin{bmatrix} x_1 & y_1 & 0 & 0 & 1 & 0 \\ 0 & 0 & x_1 & y_1 & 0 & 1 \\ x_2 & y_2 & 0 & 0 & 1 & 0 \\ 0 & 0 & x_2 & y_2 & 0 & 1 \\ x_3 & y_3 & 0 & 0 & 1 & 0 \\ 0 & 0 & x_3 & y_3 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} = \begin{bmatrix} x'_1 \\ y'_1 \\ x'_2 \\ y'_2 \\ x'_3 \\ y'_3 \end{bmatrix} \quad (1)$$

식(1)을 이용해 검사지를 정렬시킨 결과는 Fig 3과 같다.

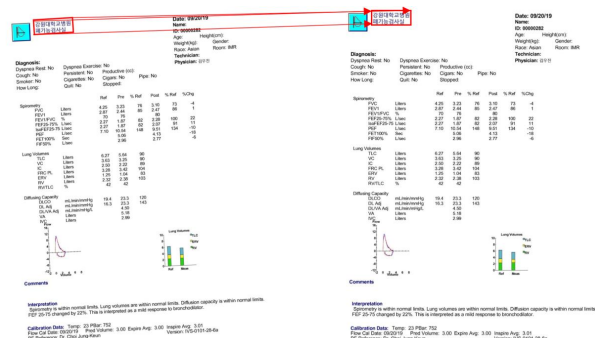


Fig. 3. Example of Image Alignment using Affine Transformation

2. Type Classification

폐 기능 검사지의 분류는 224×224 크기의 검사지 영상을 모델의 입력으로 받아서 분류를 진행한다. 실험결과에서 비교 모델군의 정확도가 유의미한 차이를 보이지 않기 때문에 검사지 분류를 위한 백본 네트워크는 파라미터가 비교적 적고 평균 FPS가 가장 빠른 ShuffleNetV2_x0.05 [10]에 분류를 위한 층(Layer)를 추가해서 사용했다. 모델 일반화와 수렴 속도 향상을 위해 ImageNet 데이터에 대해 사전학습된 가중치를 사용했으며, 손실 함수는 다중 분류 학습을 위해 Binary Cross-Entropy(BCE) Loss를 사용했다.

3. Text Extracting

글자 추출은 문자 검출과 문자 인식 두 가지 과정을 통해서 진행된다. 본 논문에서는 보다 정확한 글자 추출을 위해 통합된 모델이 아닌 문자 검출과 문자 인식을 각각 분리해서 진행하였다. 문자 검출에는 Clova AI Research의 CRAFT[11] 모델을 사용하였다. 검출 모델을 이용해 추출한 문자의 좌표정보와 이미지 분류 모델을 통해 얻어낸 검사지의 종류를 기반으로 Fig 3과 같이 이미지를 정렬한다. 문자 인식 모델에서는 정렬된 글자의 이미지를 이용해 문자를 인식한다. 문자 인식에는 CSTR[12] 모델을 사용하였다. 글자 검출과 인식 모델의 경우 모델 일반화와 수렴 속도의 향상을 위해 MJSynth와 SynthText 데이터셋에 대해 사전학습된 가중치를 사용하여 학습을 진행하였다.

4. Post processing

후처리 알고리즘은 오검출된 글자의 분리와 오인식된 글자의 보정을 위한 편집 거리 알고리즘 두 가지로 구성된다. 오검출된 글자의 분리는 Fig 4의 (a)와 같이 글자 간격이 좁아 잘못 검출된 글자를 분리하기 위해 사용한다.

Spirometry (BTPS)

FVC	Liters
FEV1	Liters
FEV1/FVC	%
FEF25-75%	L/sec
IsoFEF25-75L/sec	
FEF75-85%	L/sec
PEF	L/sec

(a) Wrong detection case

Spirometry (BTPS)

FVC	Liters
FEV1	Liters
FEV1/FVC	%
FEF25-75%	L/sec
IsoFEF25-75L/sec	
FEF75-85%	L/sec
PEF	L/sec

(b) Correct detection case

Fig. 4. Example of two detection cases

글자의 분리는 항목이나 수치가 정상적으로 검출되지 않았을 경우에 진행한다. 인식된 글자를 대상으로 유효성을 검토한 후 비정상적인 길이 혹은 항목이 발견되면 사전에 정의된 항목과의 문자열 매칭을 통해 글자를 분리한다. 오인식된 글자의 보정에는 편집 거리 알고리즘을 사용한다. 편집 거리 알고리즘은 문자열간의 유사도를 비교하는 알고리즘이다. Table1과 Table2는 오인식된 글자와 사전에 정의된 항목간의 유사도를 측정하는 과정을 보여준다.

Table 1. Similarity between 'L1ter5' and 'Liters'

	L	I	t	e	r	s
L	0	1	2	3	4	5
1	1	1	2	3	4	5
t	2	2	1	2	3	4
e	3	3	2	1	2	3
r	4	4	3	2	1	2
5	5	5	4	3	2	2

Table 2. Similarity between 'L1ter5' and 'L/sec'

	L	/	s	e	c
L	0	1	2	3	4
1	1	1	2	3	4
t	2	2	2	3	4
e	3	3	3	2	3
r	4	4	4	3	3
5	5	5	5	4	4

Table1의 경우 'L1ter5'라는 글자를 'Liters'로 변환하는데 총 2번의 대치라는 비용이 필요하고 Table2의 경우 'L1ter5'를 'L/sec'으로 변환하는데 1번의 삭제, 3번의 대치라는 비용이 필요하다. 이러한 방식으로 오인식된 글자와 사전에 정의된 항목간의 유사도와 비용을 계산하여 비용이 가장 적게 들어가는 방향으로 글자를 보정한다. 이 과정을 통해서 얻어낸 글자와 경계상자의 좌표를 이용하여 항목과 수치를 매칭한다. Fig 5은 엑셀 형태로 정형화된 데이터의 예시를 보여준다.

IV. Experiment

1. Comparison of Type Classification Models

실험에서 이미지 분류 모델의 비교군의 정확도 차이가 유의미하지 않아 이미지 분류 모델의 선정을 위한 비교 모델군의 구성은 처리속도가 빠르고 파라미터 수가 적은 모델 위주로 구성하였다. 이미지 분류 모델의 학습 및 검증에 위한 데이터셋은 학습 이미지 2,743장과 검증 이미지 914장으로 구성했다. 모델의 선정 기준은 정확도, 실행시

PBar	Height	Weight	FVC_Unit	FVC_Ref	FVC_Pre	FVC_Pre_%Ref	FVC_Post	FVC_Post_%Ref
747	149	60.0	liters	2.55	2.07	81	2.17	85
762	171	79.0	liters	4.72	3.97	84	3.90	83
744	150	60.0	liters	3.10	1.06	34	1.06	34
765	153	47.0	liters	3.28	2.31	70	2.27	69
756	171	70.0	liters	3.94	4.25	108	4.25	108

Fig. 5. Example of Structured Data

간, 파라미터를 사용했다. 평가 이미지 914장에 대한 평가 결과는 Table3와 같다.

Table 3. Results of Classification Model Test

Model	Accuracy	Params(M)	FPS
SqueezeNet1_1[13]	100.0	1.24	101.4
ShuffleNetV2_x0_5	100.0	1.37	104.3
DenseNet-121[14]	100.0	7.98	92.6
MobileNetV3-Small[15]	100.0	1.77	104.2
EfficientNet-B0[16]	100.0	5.3	96.6

Table3의 결과를 참고하여 검사지 타입 분류를 위한 모델은 ShuffleNetV2_x0_5로 선정하였다.

2. Comparison of Text Detection and Recognition Models

문자 검출 모델과 문자 인식 모델의 선정을 위한 비교 모델군의 구성은 ICDAR2013과 ICDAR2015 데이터셋에 대해서 우수한 성능을 보이는 모델을 위주로 구성하였다. 문자 검출 모델의 비교를 위한 평가지표는 실행시간과 F1 점수를 참고하였으며, 검사지에서 검출하지 못하는 글자가 있는 모델은 비교대상에서 제외하였다. 문자 인식 모델의 비교를 위한 평가지표는 의료 데이터의 특성상 인식이 잘 못될 경우에는 환자에 대한 해석에 큰 문제를 일으키므로 실행시간을 배제하고 인식된 문자의 정확도만을 비교한다.

문자 검출 모델과 문자 인식 모델의 학습 및 평가 데이터셋은 각각 1,520장의 학습 이미지, 340장의 검증 이미지, 340장의 평가 이미지로 구성하였다. 검사지 당 평균적으로 120개의 단어가 사용됐으며, 총 262,844개의 단어를 사용하였다. 문자 검출 모델의 평가결과는 Table4, 문자 인식 모델의 평가결과는 Table5와 같다.

Table 4. Results of Text Detection Models Test

Method	R	P	F1	FPS
TextFuseNet[17]	97.1	99.2	98.1	1.1
CRAFT	97.2	98.9	98.0	2.4
DB	94.4	97.3	95.8	1.8
Mask TextSpotter	95.1	97.6	96.3	1.2
SBD	96.6	98.9	97.7	1.1

Table 5. Results of Text Recognition Models Test

Method	Accuracy
CSTR	98.6
RCEED[18]	96.7
SATRN[19]	96.3
Text Scanner[20]	95.8
Yet Another Text Recognizer	98.4

위 결과를 참고하여 문자 검출 모델은 Clova Research의 CRAFT로, 문자 인식 모델은 CSTR로 선정하였다. 실험 환경은 Table6과 같다.

Table 6. Experiment environment

OS	Ubuntu 18.04
CPU	Ryzen 5 5600X-6
GPU	RTX 3060
RAM	32GB
Framework	pytorch

3. System Architecture

최종적으로 제안하는 시스템의 흐름은 다음과 같다. 입력받은 이미지를 224×224 사이즈로 조절한 후에 검사지의 타입 분류를 위한 네트워크인 ShuffleNetV2_x0_5를 통과시켜 검사지 타입을 획득한다. 이후 CRAFT 모델을 이용해 검사지의 문자를 추출하고 보다 정확한 글자 인식을 위해 추출된 글자와 검사지의 타입을 통해 왜곡된 이미지를 정렬한다. 정렬된 이미지는 CSTR 모델을 통과해 글자로 변환된다. 위 과정을 통해 얻어낸 글자에 대한 정확도 향상을 위해 제안한 후처리 과정을 추가로 진행한다. Table7은 후처리 과정을 적용하기 전과 후의 정확도 차이를 보여준다.

Table 7. The difference in accuracy before and after applying the post-processing process.

Method	Accuracy	Throughput (Image/sec)
CRAFT + CSTR	98.6	1.96
CRAFT + CSTR + Separate	98.8	1.99
CRAFT + CSTR + Separate + Edit distance	99.2	2.05

마지막으로 보정된 글자와 글자의 좌표 값을 기반으로 항목과 수치를 매칭하여 정형화한다. 제안하는 시스템은 평균적으로 검사지 한 장을 처리하는데 2.0초 내외의 실행 시간을 보이며, 99.2%의 정확도를 보여준다.

V. Conclusion

본 논문에서는 딥러닝 모델을 이용한 폐 기능 검사지의 정형화 시스템을 제안하였다. 딥러닝 모델은 이미지 분류 네트워크, 문자 검출 모델, 문자 분류 모델을 사용했으며, 제안한 전처리와 후처리 과정을 통해 시스템의 정확도를 향상했다. 제안하는 시스템은 검사지 한 장 당 2초 내외의 실행시간과 99.2%의 높은 정확도를 얻어냈다. 해당 결과는 전문화된 인력의 수작업 대비 1/60 수준의 실행시간과 약 4% 높은 정확도이다. 하지만 제안한 시스템은 딥러닝 기술의 특성으로 인하여 학습되지 않은 검사지와 단어에 대해 정형화를 진행할 수 없는 문제점이 있다. 차후 진행 될 연구에서는 다양한 종류의 검사지와 단어에 대한 학습을 통해 이러한 문제점을 극복할 예정이다.

ACKNOWLEDGEMENT

This research was supported by National Information Society Agency (NIA) funded by the Ministry of Science, ICT through the Big Data Platform and Center Construction Project (No. 2020-Data-W123).

REFERENCES

- [1] Max Jaderberg, Karen Simonyan, Andrea Vedaldi, and Andrew Zisserman. "Reading Text in the Wild with Convolutional Neural Networks" *Int J Comput Vis* 116, 1-20 May. 2015. DOI: <https://doi.org/10.1007/s11263-015-0823-z>
- [2] Ankush Gupta, Andrea Vedaldi, and Andrew Zisserman. "Synthetic Data for Text Localisation in Natural Images" *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2315-2324 Apr. 2016. DOI: 10.1109/CVPR.2016.254
- [3] Xing, Linjie, et al. "Convolutional Character Networks" *IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 9126-9136 2019. DOI: 10.1109/ICCV.2019.00922
- [4] Liu, Yuliang, et al. "Exploring the Capacity of an Orderless Box Discretization Network for Multi-orientation Scene Text Detection" *Int J Comput Vis* 129, 1972-1992 2021. DOI: <https://doi.org/10.1007/s11263-021-01459-7>
- [5] Lyu, Pengyuan, et al. "Mask TextSpotter: An End-to-End Trainable Neural Network for Spotting Text with Arbitrary Shapes" *European Conference on Computer Vision (ECCV)*, pp. 71-88 Aug. 2018. DOI: https://doi.org/10.1007/978-3-030-01264-9_5
- [6] Liao, Minghui, et al. "Real-time Scene Text Detection with Differentiable Binarization" *AAAI Conference on Artificial Intelligence*, pp. 11474-11481 April. 2020. DOI: 10.1609/aaai.v34i07.6812
- [7] Liu, Jingchao, et al. "Pyramid mask text detector." *arXiv preprint arXiv:1903.11800* Mar. 2019.
- [8] Huang, Zhida, et al. "Mask R-CNN with pyramid attention network for scene text detection." *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 764-772 IEEE, 2019. DOI: 10.1109/WACV.2019.00086
- [9] Loginov, Vladimir. "Why You Should Try the Real Data for the Scene Text Recognition." *arXiv preprint arXiv:2107.13938* Jul. 2021.
- [10] Zhang, Xiangyu, et al. "Shufflenet: An extremely efficient convolutional neural network for mobile devices." *IEEE conference on computer vision and pattern recognition (CVPR)*, pp. 6848-6856 June. 2018. DOI: 10.1109/CVPR.2018.00716
- [11] Baek, Youngmin, et al. "Character region awareness for text detection." *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 9365-9374 Apr. 2019. DOI: 10.1109/CVPR.2019.00959
- [12] Cai, Hongxiang, et al. "Revisiting Classification Perspective on Scene Text Recognition." *arXiv preprint arXiv:2102.10884* Jun. 2021.
- [13] Iandola, Forrest N., et al. "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size." *arXiv preprint arXiv:1602.07360* Dec. 2017.
- [14] Huang, Gao, et al. "Densely connected convolutional networks." *IEEE conference on computer vision and pattern recognition (CVPR)*, pp. 4700-4708 Jan. 2018. DOI: 10.1109/CVPR.2017.243
- [15] Howard, Andrew, et al. "Searching for mobilenetv3." *IEEE/CVF International Conference on Computer Vision*, pp. 1314-1324 Nov. 2019. DOI: 10.1109/ICCV.2019.00140
- [16] Tan, Mingxing, and Quoc Le. "Efficientnet: Rethinking model scaling for convolutional neural networks." *International Conference on Machine Learning (PMLR)*, pp. 97:6105-6114 May. 2019.
- [17] Ye, Jian, et al. "TextFuseNet: Scene Text Detection with Richer Fused Features." *IJCAI*, pp. 516-522 July. 2020. DOI: 10.24963/ijcai.2020/72

- [18] Cui, Mengmeng, et al. "Representation and Correlation Enhanced Encoder-Decoder Framework for Scene Text Recognition." arXiv preprint arXiv:2106.06960 Jun. 2021.
- [19] Lee, Junyeop, et al. "On recognizing texts of arbitrary shapes with 2D self-attention." IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 546-547 2020. DOI:10.1109/CVPRW50498.2020.00281
- [20] Wan, Zhaoyi, et al. "Textscanner: Reading characters in order for robust scene text recognition." AAAI Conference on Artificial Intelligence. 34,07 pp. 12120-12127 Jan. 2020. DOI:10.1609/AAI.V34I07.6891



Sung-Ok Kwon obtained a Masters and Doctorate in Food and Nutrition from Kyung Hee University, Korea, in 2003 and 2013, respectively. Dr. Kwon is currently research professor of the Interdisciplinary Graduate

Program in Medical Bigdata Convergence at Kangwon National University, Chunchen-si, Korea, in 2021. My primary research efforts focus on environmental diseases, dietary exposure assessment, nutritional epidemiology, and biostatistics.



Woo-Jin Kim received a B.S degree in 1994, an M.S. degree in 2004, in Medical College from Seoul National University, and a PhD. degree in 2006 in Medicine from Hallym university.

In 2004, he joined the Department of Internal Medicine Kangwon National University, where he is currently a professor. His research interests are in the areas of biomedical informatics.

Authors



Sang-Hyun Jo received the B.S. degrees in Computer Science and Engineering from Kangwon National University, Korea, in 2020. He is currently a M.S. Student in the department of Computer Science and

Engineering at Kangwon National University. His research interests are in the areas of machine learning and computer vision.



Dae-Hoon Kim received the B.S. degrees in Computer Science and Engineering from Kangwon National University, Korea, in 2020. He is currently a M.S. Student in the department of Computer Science and

Engineering at Kangwon National University. His research interests are in the areas of machine learning and computer vision.



Yoon Kim received a B.S. degree in 1993, an M.S. degree in 1995, and a Ph.D. degree in 2003, in electronic engineering with the Department of Electronic Engineering from Korea University In 2004, he joined the

Department of Computer Science and Engineering, Kangwon National University, where he is currently a professor. From 1995 to 1999, he was with the LG-Philips LCD Co., where he was involved in research and development on digital image equipment. His research interests are in the areas of machine learning, multimedia communications, and computer vision.



Sang-Ah Lee obtained a Masters and Doctorate in Food and Nutrition from Seoul National University, Korea, in 1996 and 2001, respectively. Dr. Lee joined the faculty of the Department of Preventive Medicine,

School of Medicine at Kangwon National University, Chunchen-si, Korea, in 2008. She is currently professor of the Department of Preventive Medicine at School of Medicine, and the chair of the Interdisciplinary Graduate Program in Medical Bigdata Convergence at Kangwon National University. My primary research efforts focus on cancer epidemiology, nutritional epidemiology, and biostatistics.