코로나19로 또래와 단절된 아동을 위한 인공지능 낙서 로봇 제안

송주연O, 이강희*
OO숭실대학교 글로벌미디어학부,
*숭실대학교 글로벌미디어학부
e-mail: superyodi.song@gmail.comO, kanghee.lee@ssu.ac.kr*

Proposal of AI-based Graffiti Robot for Children disconnected from Peers with COVID-19

Ju-Yeon SongO, Kang-Hee Lee*
OGlobal School of Media, Soongsil University,
*Global School of Media, Soongsil University

● 요 약 ●

본 논문에서는 코로나19 사태로 인한 팬데믹(pandemic) 현상으로 인해 또래와 단절된 아동들의 정서발달을 위해 인공지능 낙서 로봇인 Doodle Robot을 제안한다. Doodle Robot은 또래 형제가 없는 아동에게 함께 그림을 그릴 수 있는 그림친구로서 아동의 정서적 발달에 기여한다. YOLO 알고리즘을 사용하여 객체검출기능을 구현하였고 낙서 Data는 Quick! Draw Dataset에서 추출하였다.

키워드: YOLO, Quick Draw, 딥러닝, 인공지능, 로봇, COVID-19

I. Introduction

코로나 19사태가 장기화함에 따라 유치원이나 초등학교에 가지 못하게 된 어린이들을 대상으로 '재택교육'과 관련된 상품에 대한 수요가 급증하면서 온라인 교육 프로그램에 대한 제품이 늘고 있다. 하지만 또래로부터 단절되어 집에 고립된 아동의 정서발달에 관련된 제품은 부족한 상황이다.

이러한 문제를 해결하기 위해 본 논문은 인공지능 낙서 로봇인 Doodle Robot을 제안한다. Doodle Robot은 사용자가 사진을 입력하면 사진 속 객체에 대응하는 낙서를 그려주는 로봇이다. Doodle Robot은 또래 형제가 없는 아동에게 함께 그림을 그릴 수 있는 그림 친구로서 아동의 정서적 발달에 이바지한다. 또한, Doodle Robot을 따라 스케치북에 크레파스로 그림을 그리는 놀이는 아날로그적 놀이로서 아동기에 해로운 영향을 미치는 스마트폰을 벗어날 수 있는 계기가 될 수 있으므로 아동 정서발달에 긍정적인 영향을 미친다[1].

II. Methodology

이 장에서는 Doodle Robot에 사용된 객체검출 알고리즘과 Dataset에 대해 설명한다.

1. 객체검출 알고리즘

사진 속 객체를 검출하기 위해서 YOLO v3-tiny를 사용하였다[2].

1.1 R-CNN 알고리즘

R-CNN 알고리즘은 Fig. 1과 같이 3개의 네트워크층으로 이루어져 있다[3]. 맨 앞단의 CNN(Convolutional Neural Network)이 객체의 특징을 추출하는 역할을 하고, 두 번째 단의 RPN(Region Proposal Network)는 객체가 있을 법한 경계박스를 추출하며 마지막 네트워크인 classifier는 앞의 두 네트워크가 전달하는 객체의 특징이 담긴 데이터와 경계박스의 위치를 받아서 분류하는 역할을 한다.

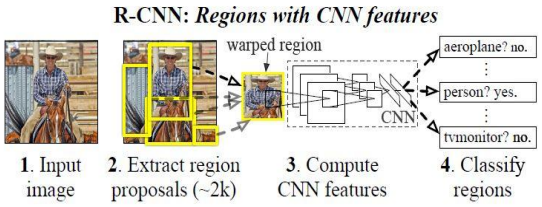

Fig. 1. R-CNN의 구조

Fig. 2에서 확인할 수 있듯이 PRN은 학습 과정에서 2000개, 테스트 과정에서는 약 300여 개의 영역을 선정하기 때문에 객체검출과

정이 복잡하고 느려서 상대적으로 과정이 간단하고 빠른 YOLO 알고리즘으로 객체검출기능 구현하고자 했다.

train-time region proposals		test-time region proposals		
method	# boxes	method	# proposals	mAP (%)
SS	2000	SS	2000	58.7
EB	2000	EB	2000	58.6
RPN+ZF, shared	2000	RPN+ZF, shared	300	**59.9**

Fig. 2. Faster R-CNN의 후보군

1.2 YOLO(You Only Look Once) 알고리즘

YOLO 알고리즘은 proposal의 수를 줄이기 위해서 Fig. 3과 같은 grid 방식을 이용해서 객체를 검출한다.

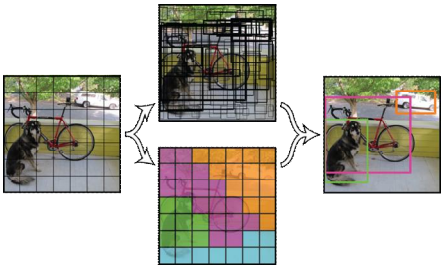

Fig. 3. grid를 이용한 YOLO

YOLO 알고리즘의 구조는 Fig. 4과 같으며 네트워크 마지막 단계에서 경계박스의 위치와 객체 분류가 동시에 이루어지는 1-layer 구조로 이루어지기 때문에 검출시간이 R-CNN보다 약 10배 이상 빠르다.

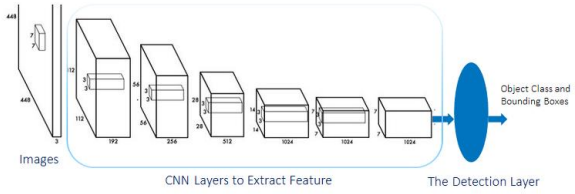

Fig. 4. YOLO의 구조

2. Dataset

2.1 Quick, Draw! Dataset

Fig. 5. Quick, Draw! Dataset[4]

사진 속 객체를 검출하기 위해서 YOLO v3-tiny를 사용하였다. 다른 객체검출 알고리즘인 R-CNN보다 더 빨라서 사용했고 YOLO v3-tiny darket을 사용하였다.

Quick, Draw! Dataset 에는 Google이 만든 Catch Mind 게임인 Quick, draw![5]의 플레이어들에 의해서 만들어진 345종류의 그림 데이터 5천만 개가 있다. 그림 데이터들은 Fig. 6과 같은 구조로 저장되어 있어 timestamp와 drawing 데이터를 이용하여 획(stroke) 단위로 그림을 재구성할 수 있다.


```
{
    "key_id":"5891796615823360",
    "word":"nose",
    "countrycode":"AE",
    "timestamp":"2017-03-01 20:41:36.70725 UTC",
    "recognized":true,

"drawing":[[[129,128,129,129,130,130,131,132,132,1
33,133,133,133,...]]]
  }
```


Fig. 6. 그림 데이터 한 개의 구조

III. The Proposed Scheme

본 로봇은 사용자가 입력한 사진에서 객체를 검출한 후 낙서 dataset에서 임의의 낙서를 그려주는 딥러닝 기반의 로봇이다. 사용자에게 낙서를 그려주기 때문에 Doodle Robot으로 이름을 붙였다.

Doodle Robot은 Python언어로 개발되었고, darknet 프레임워크의 YOLO v3-tiny.cfg를 사용해서 객체검출 기능을 구현하였다[6]. 로봇이 그리는 낙서는 Quick, Draw! Dataset를 사용하였고, quickdraw 라이브러리를 이용해서 Python 언어로 Dataset을 불러올 수 있었다[7]. 로봇은 Dobot을 사용하였다.

Fig. 7은 낙서 로봇 프로그램에 대한 순서도이다.

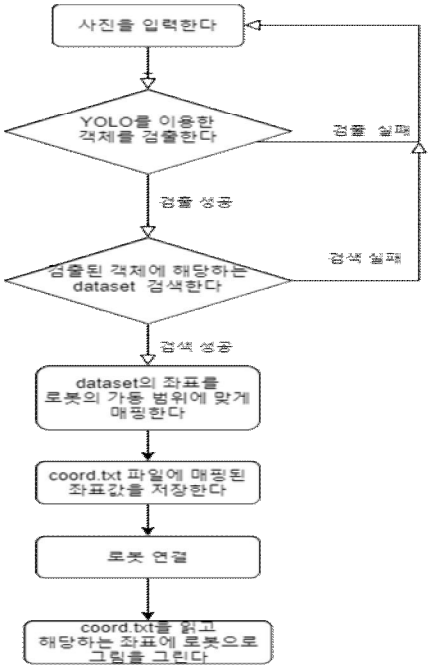

Fig. 7. 낙서 로봇 프로그램의 순서도

사용자가 사진을 입력하고 낙서 로봇 프로그램을 실행하면 YOLO 알고리즘으로 사진 속 객체를 검출해서 Fig. 8과 같이 경계박스를 그리고, 라벨(객체이름)을 표시한다.

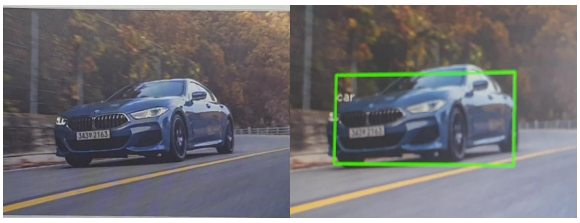

Fig. 8. 사용자가 입력한 사진과 검출된 객체

Quick, Draw! Dataset 의 갈래에 label이, Fig 8에 따르면 label은 car가 된다, 존재하면 quick draw 라이브러리를 사용하여 Google Dataset 서버에서 label에 대한 모든 data를 내려받는다.

Fig. 9은 최종적으로 낙서 로봇이 Fig. 8에 해당되는 낙서를 수행완료한 것을 볼 수 있다. 좀더 자세한 연구와 결과는 추후 본 논문의 확장판에서 기술하도록 한다.

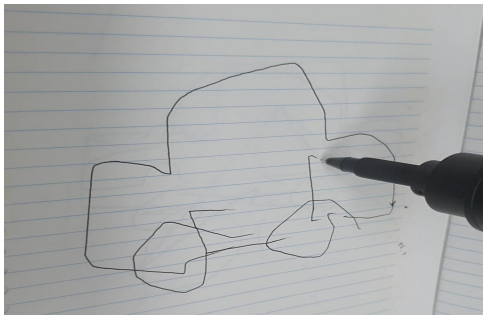

Fig. 9. 낙서 로봇

IV. Conclusions

인공지능 낙서 로봇인 Doodle Robot은 디지털미디어에 노출된 아동에게 종이와 펜을 이용한 그림 그리기 활동과 같은 아날로그적 경험을 선사하고 로봇과 함께 그림을 그리는 과정에서 아동이 로봇에게 정서적인 교감을 느낄 수 있다는 점에서 의의가 있다.

더 나아가 Doodle Robot에 스마트폰을 연동시켜 스마트폰으로 사진을 찍으면 바로 로봇이 그림을 그리게 만들고, 로봇의 외형을 아동들이 좋아하는 캐릭터로 꾸민다면 고급 아동용 로봇으로서 상용화 될 수 있을 것이다.

ACKNOWLEDGEMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음(2018-0-00209)

REFERENCES

[1] Ji Eun Kim. "The effects of smartphone dependency traits on the peer interactions of preschool children" VOL.-NO.- (2018)

[2] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, “The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788

[3] Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks.“ Advances in neural information processing systems. 2015.

[4] The Quick, Draw! Dataset, https://github.com/googlecreativelab/quickdraw-dataset

[5] QUICK, DRAW!, *https://quickdraw.withgoogle.com*

[6] Redmon, Joseph, and Ali Farhadi. "Yolov3: An incrementa improvement.“ arXiv preprint arXiv:1804.02767, 2018.

[7] Quickdraw_python Library, *https://github.com/martinohanlon/quickdraw_python*