

영어 교육용 콘텐츠를 제공하는 로봇 설계 및 구현

김민지*, 김가은*, 이상원**

*서경대학교 컴퓨터과학과

**송실대학교 전자정보공학부 IT융합과

jee00609@naver.com, rkdms7220@naver.com, lee2155507@naver.com

Design and implementation of Robot providing English educational contents

Min-Jee Kim*, Ga-eun Kim*, Sang-won Lee**

*Dept. of Computer Science, Seo-Kyeong University

**Dept. of Electronic Engineering, Soong-sil University

요 약

코로나바이러스 감염증-19(COVID-19)로 인해 기존 대면 수업 위주의 학습 방식에서 비대면 온라인 수업으로 변화하고 있다. 양육자가 일터에 나가야 하는 취약계층에서는 온라인 학습 여건을 갖추지 못해 학습 부진 현상이 발생하고 있다. 이에 본 논문에서는 가정 내에서 자녀들이 영어 학습을 돕는 로봇을 통해 공부에 흥미를 느낄 수 있기를 기대한다.

1. 서론

코로나바이러스 감염증-19가 확산함에 따라 정부에서는 감염병 예방에 관한 법률 제50조에 따라 전국 유·초·중·고 신학기 개학을 연기했다. 이후 사태가 지속하면서 4차에 걸친 학교 휴업 명령 끝에 4월 9일부터 단계적 온라인 개학을 시행했다. 원격 수업이 본격화됨에 따라 중위권 학생들의 성적이 하락하는 ‘학습 격차’ 현상이 전국에서 발생하고 있다.[1] 특히 학교에서 담당했던 돌봄 기능의 공백에 대하여 초등학생 자녀를 둔 맞벌이 학부모 혹은 경제적 상황이 어려운 학부모들이 더 부담스럽게 느끼고 있다.[2]

경제적 수준이 낮은 가정의 학생은 상대적으로 경제적 수준이 높은 가정의 학생보다 온라인 수업의 진도를 따라가는 데 어려움을 느끼고 있다. 이는 경제적 수준이 낮을수록 학생의 전용 디지털기기 보유율이 현저히 낮고, 온라인 수업에 집중하기 어려운 장소에서 공부해야 하는 상황 속에서 온라인 수업에 부정적이다.[3]

이에 본 논문에서는 초등학교의 대표적인 과목 국·영·수 과목 중 영어 과목을 대상으로 학습을 도우면서 아이의 안전을 지키는 침입자 알림 기능을 결합한 로봇을 설계하고 이를 구현하였다. 본 논문에서 설계하고 구현한 로봇은 3D 프린터를 통해 출력한 조립형 하드웨어와 오픈소스 하드웨어 라즈베리파이를 결합한 로봇으로, 가정 내에서 저학년 학생이 쉽

게 조작하여 혼자서도 공부할 수 있도록 돕는 ‘Turing Robot’을 제작하였다. 본 논문에서 설계한 로봇의 주요 기능은 영어 교육 콘텐츠와 더불어 자녀 안심 서비스 중 IOT 기술인 침입자 알림 기능(홈 CCTV기능)이 존재한다. 또한 터치 디스플레이를 통해 손의 터치만으로 모든 기능을 사용할 수 있도록 제작하여 사용자의 편의성을 높였다. 본 논문에서 설계하고 구현한 로봇은 경제 수준이 어려운 가정에서도 사용 혹은 제작할 수 있도록 저가이면서도 저학년생들이 편하게 이용할 수 있도록 간단한 내구성을 추구한다. 따라서 구글에서 제공하는 오픈소스 데이터셋과 정부에서 무료로 사용할 수 있도록 지원하는 3D 프린팅 기술[4] 등 저가이면서도 쉽게 박살 나지 않는 재질을 이용하여 설계하였다. 또한 모든 콘텐츠의 코드와 3D 모델링을 오픈소스화하여 사용자들이 자체적으로 커스텀 하여 제작할 수 있는 교육 로봇으로 사용되리라 기대한다.

2. 사용 기술 및 선행 연구 사례

2.1 온라인 필기 인식 (Multi-Language Online Handwriting Recognition)

온라인 필기 인식 기술은 자유형 필기 입력에서 구조화된 패턴을 인식하는 것으로 구성된다. 인간-컴퓨터 상호 작용으로서의 연구를 통해 모바일 터치 상황에서 언어에 대한 인식기를 구축한다. 이 기술을 바탕으로 사람의 그림 데이터 셋을 모아 기계 학

습 분류기를 구축하였다. 본 논문의 교육 콘텐츠 중 하나인 QuickDraw는 사용자의 낙서를 인식하여 영어 단어를 유추해내는 기능으로 필기 인식 정보 제공을 기반 기술로 사용하고 있다.

2.2 광학 문자 인식 (Optical Character Recognition)

광학 문자 인식은 인쇄물이나 사람이 쓴 문자의 영상을 기계가 읽을 수 있는 문자로 변환하는 기술이다. 본 논문의 교육 콘텐츠 중 하나인 한-영 번역기는 AWS Textract의 OCR 기능을 결합하여 만든 기술이다. 사용자가 번역하고자 하는 사진을 안드로이드 애플리케이션으로 전송하면 로봇이 이미지 속의 텍스트들을 추출하여 번역할 수 있도록 하였다.

2.3 웹 크롤링 (Web Crawling)

웹 크롤링은 웹상의 html로 작성된 방대한 정보를 가져오는 기술이다. 본 논문의 교육용 콘텐츠 중 하나인 'GUI를 이용한 Voice&text'에서 이용한다. 어린이의 흥미를 끌 수 있는 어린이 영어 뉴스와 동화를 위주로 텍스트를 가져와 txt 형태로 파일 스토리지(Amazon S3)에 저장하고 사용자가 접근할 수 있도록 한다.

2.4 안면 인식 기술(Facial recognition technology)

Facial recognition technology는 선택된 얼굴 특징과 안면 데이터베이스를 서로 비교하는 기술로 클라우드 서비스인 Amazon Rekognition을 이용한다. Amazon Rekognition은 딥 러닝 기술을 사용하여 애플리케이션에 이미지 및 비디오 분석을 쉽게 할 수 있는 기술이다. 본 논문에서는 Raspberry Pi와 Amazon Rekognition을 이용하여 안면 데이터베이스 이미지와 Raspberry Pi로 들어온 안면 이미지를 비교하는 기능을 수행한다.

2.5 음성 합성 시스템(Text to speech)

음성 합성 시스템은 텍스트를 음성으로 바꿔주는 기술이다. 본 논문에서는 'Voice&text output'과 '한-영 번역 OCR 과 Textract'에서 사용하는 기술로 클라우드 서비스인 Amazon Polly를 이용하여 기능을 수행한다.

2.6 서버리스 플랫폼 (Serverless Platform)

서버리스 플랫폼이란 서버를 고려하지 않고 애플리

케이션을 구축하고 실행할 수 있는 아키텍처를 말한다.

서버 또는 클러스터 프로비저닝, 패치 적용, 운영체제 유지 관리 및 용량 프로비저닝과 같은 인프라 관리 작업을 덜어낸다.

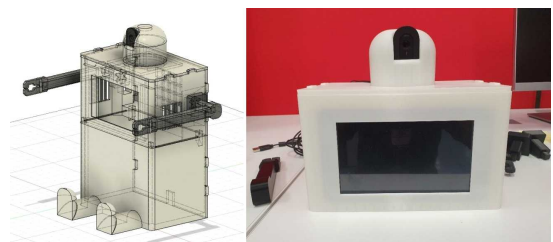
거의 모든 유형의 애플리케이션 또는 백엔드 서비스를 서버리스로 구축할 수 있으며, 2.2 광학문자 인식을 이용한 서비스와, 2.4 안면 인식 기술을 이용한 서비스가 서버리스 플랫폼을 이용하여 구현되었다.

2.7 선행 연구 사례 (Web Crawling)

교육용 로봇의 선행 사례로는 AKA에 의해 개발된 인공지능(AI) 대화엔진을 가진 소셜 로봇 'Musio'[5]가 존재한다. 현재 제주도 무릉 초·중학교에서 도입하였으며 교과서 영어 회화 연습 등을 제공하고 9만 8000엔(한화 약 110만원)으로 구매 가능하다.

3. 로봇 설계 및 사용 시나리오

3.1 로봇 설계

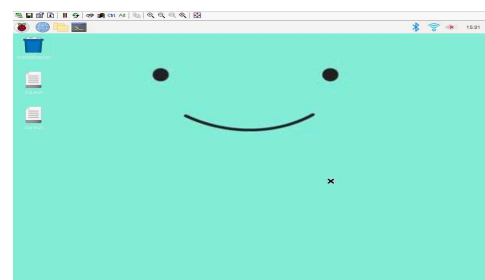


[그림1]로봇 디자인과 실제 출력물

그림 1은 123D Design을 통해 모델링한 로봇의 디자인과 실제로 출력한 로봇의 모습이다. 로봇의 머리에 카메라를 부착하고 몸 앞에 터치 디스플레이를 삽입할 수 있도록 제작하였다. 사용자의 입력에 반응할 수 있도록 모터가 부착된 팔로 설계하였다.

3.2 사용 시나리오

3.2.1 초기 화면 및 침입자 알림 기능



[그림2] 터치 디스플레이 화면

그림2의 사진 속 터치 디스플레이의 바탕화면에서 원하는 영어 교육 콘텐츠 아이콘을 클릭하면 원하는 기능이 실행된다. 침입자 알림 기능은 백그라운드에서 실행되도록 구현하여 다른 기능을 사용하고 있어도 계속해서 외부인의 침입을 감지할 수 있다.

3.2.2 영어 교육 콘텐츠 화면



[그림3] QuickDraw 및 Reading 서비스 사용 화면

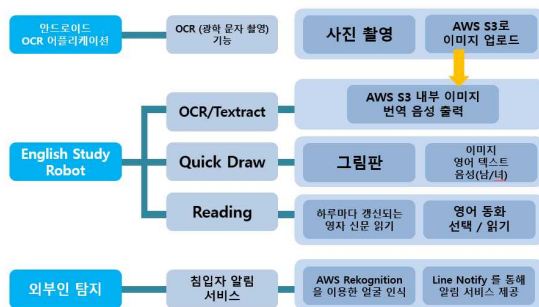
그림3의 화면들은 각각 로봇에서 제공하는 영어 교육용 콘텐츠다. QuickDraw 및 영어 Reading 기능이다. QuickDraw 기능은 사용자가 흰색 바탕에 그림을 그리고 Predict 버튼을 클릭하면 로봇은 낙서와 가장 유사한 대상을 추론하고 이미지, 텍스트와 함께 음성으로 설명해준다.

Reading 기능을 클릭 시 하루마다 갱신하는 영어 신문의 기사 중 하나를 클릭할 수 있다. 원하는 기사를 클릭하면 기사와 함께 영어 음성을 출력한다.

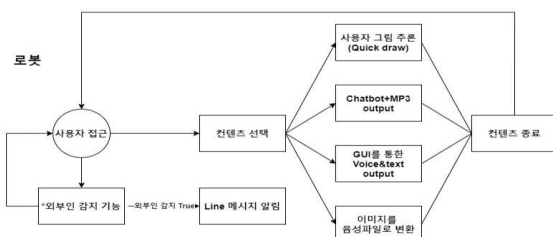
4. 설계 및 구현

4.1 시스템 구조

전체 시스템의 구조와 흐름도는 그림4와 그림5와 같다.



[그림4] 전체 시스템의 구조



[그림5] 전체 시스템의 흐름도

4.2 콘텐츠 구현 알고리즘

4.2.1 QuickDraw, QuickDraw Python 알고리즘

```

20 def predict():
21     filename = "saveImage/image.png"
22     image = save(filename)
23     model = load_model("QuickDraw.h5")
24     file = "saveImage/image.png"
25     img = cv2.imread(file)
26     img = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
27     img = cv2.resize(img, (28, 28))
28     img = img.reshape((1, 1, 1, 28, 28))
29     pred = model.predict(img)
30     pred = np.argmax(pred, axis=-1)
31     return pred
32
33 if __name__ == '__main__':
34     img = cv2.imread("image.png")
35     img = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
36     img = cv2.resize(img, (28, 28))
37     img = img.reshape((1, 1, 1, 28, 28))
38     pred = model.predict(img)
39     pred = np.argmax(pred, axis=-1)
40     return pred
41
42 if __name__ == '__main__':
43     img = cv2.imread("image.png")
44     img = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
45     img = cv2.resize(img, (28, 28))
46     img = img.reshape((1, 1, 1, 28, 28))
47     pred = model.predict(img)
48     pred = np.argmax(pred, axis=-1)
49     return pred
50
51 if __name__ == '__main__':
52     img = cv2.imread("image.png")
53     img = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
54     img = cv2.resize(img, (28, 28))
55     img = img.reshape((1, 1, 1, 28, 28))
56     pred = model.predict(img)
57     pred = np.argmax(pred, axis=-1)
58     return pred
59
60 if __name__ == '__main__':
61     img = cv2.imread("image.png")
62     img = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
63     img = cv2.resize(img, (28, 28))
64     img = img.reshape((1, 1, 1, 28, 28))
65     pred = model.predict(img)
66     pred = np.argmax(pred, axis=-1)
67     return pred
68
69 if __name__ == '__main__':
70     img = cv2.imread("image.png")
71     img = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
72     img = cv2.resize(img, (28, 28))
73     img = img.reshape((1, 1, 1, 28, 28))
74     pred = model.predict(img)
75     pred = np.argmax(pred, axis=-1)
76     return pred

```

[그림6] 사용자 그림 인식 코드

로봇이 사람의 낙서를 인식할 수 있도록 Numpy로 구성된 데이터셋들[6]을 신경망 알고리즘인 CNN 알고리즘을 이용해 학습 모델을 제작한다.

사용자가 그린 이미지를 입력받고 해당 이미지를 OpenCV[7]를 통한 전처리 과정을 거치게 하여 사용자의 입력값을 제외한 변수를 제거한다.

전처리 과정을 거친 이미지 파일에 대해 그림6과 같이 이미지의 대상을 추론하여 값을 출력하고 로봇의 모터가 작동한다.

4.2.2 한-영 번역, Textract

사용자는 안드로이드 애플리케이션을 통해 이미지를 S3에 전송한다. 해당 이미지에 있는 내용은 OCR 기능을 이용한 Textract 서비스 코드를 적용하여 한글로 번역 후 MP3 File로 변환하여 사용자에게 들려준다.

4.2.3 Reading, GUI를 이용한 Voice&text output



[그림7] GUI 화면

웹 크롤링으로 받아온 텍스트를 Amazon Polly(TTS 기능)를 이용하여 음성파일로 변환한다. 사용자는 원하는 카테고리의 파일을 선택할 수 있다. 선택 시 변환된 음성파일과 텍스트를 그림7의 GUI화면과 같이 동시에 보여준다.

4.2.4 침입자 알림 기능, AWS Rekognition

```
while True:
    ret, img = cam.read()
    img = cv2.flip(img, 1) # Flip vertically
    gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)

    faces = faceCascade.detectMultiScale(
        gray,
        scaleFactor = 1.2,
        minNeighbors = 5,
        minSize = (int(minW), int(minH)),
    )

    for (x,y,w,h) in faces:
        cv2.rectangle(img, (x,y), (x+w,y+h), (0,255,0), 2)
        id, confidence = recognizer.predict(gray[y:y+h,x:x+w])
        # Check if confidence is less than 100 => "0" is perfect match
        if (confidence < 100):
            print(confidence)
            id = names[id]
            cv2.imwrite("image/user-"+id+".jpg", gray[y:y+h,x:x+w])
            os.system("python line.py")
            res=requests.post(url,headers=headers,files=msg)
        else:
            id = "unknown"
            confidence = " " + "{0}%".format(round(100 - confidence))

            cv2.putText(img, str(id), (x+5,y-5), font, 1, (255,255,255), 2)
            cv2.putText(img, str(confidence), (x+5,y+h-5), font, 1, (255,255,0), 1)

    cv2.imshow('camera',img)
    k = cv2.waitKey(10) & 0xFF # Press 'ESC' for exiting video
    if k == 27:
        break
# Do a bit of cleanup
print("\n [INFO] Exiting Program and Cleanup stuff")
```

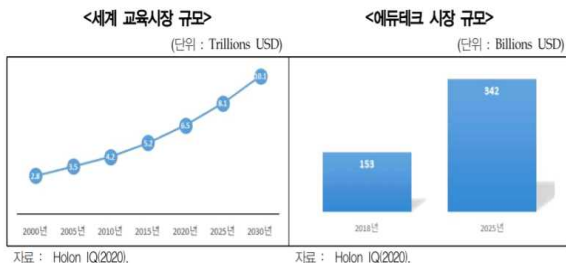
[그림8] 사용자 구성원과 외부인을 구별하는 코드

AWS Rekognition[8] 과 Line Notify를 결합하여 만든다. 외부의 상황을 알기 위해 라즈베리 파이에서 실시간으로 사진을 찍고 Amazon S3에 저장한다. 이후 Lambda 서버리스 S3 트리거를 통해 그림7과 같이 구성원과 외부인을 구별하는 Rekognition 기능을 실행한다.

이후 외부인일시 Line Notify API를 사용하여 사용자에게 알림 메시지를 전송한다.

5. 시장 조사

5.1. 에듀테크 시장 조사



[그림9]세계 교육시장 규모와 에듀테크 시장 규모

에듀테크(EduTech, Education+Technology)는 교육 분야에 ICT 기술을 융합한 새로운 교육 기술 혹은 서비스를 의미한다. 전 세계 교육 시장 규모와 에듀테크 시장 규모는 그림 9와 같이 지속적인 성장세를 보이며 저가형으로 설계된 본 논문의 로봇은 충분한 가격 경쟁력을 가지고 있다.[9]

6. 결론

코로나바이러스 감염증-19 이후 기존의 대면 수업에서 제공할 수 있었던 상호작용 기반의 수업 형식이 불가한 시대가 왔다. 온라인 학습으로 단편적인

지식과 정보를 얻는 학습에는 분명한 한계가 있다. 가정에서 충분한 교육 돌봄을 받지 못하는 저학년 학생들을 위해 상호작용이 가능한 영어 교육 로봇을 개발하였다.

비대면 교육에서 제공하지 못하는 현장감과 상호 의사소통을 제공하기 위해 사용자의 입력에 반응하는 영어 교육 로봇을 통해 경제적으로 사교육을 받는데 부담되는 가정의 자녀에게 도움이 될 것으로 예측한다. 또한 맞벌이 가정의 학부모 또한 침입자 알림 기능을 쉽게 이용하면서 혼자 집에 있을 자녀들의 안전에 대한 걱정을 줄일 수 있다고 기대한다.

[본 논문은 과학기술정보통신부 정보통신창의인재 양성사업의 지원을 통해 수행한 ICT멘토링 프로젝트 결과물입니다.]

참고문헌

- [1] 전국교직원노동조합, 보도자료 "코로나19 상황, 2020년 1학기 교육실태와 교사 요구조사 결과", P.4, 2020.08.21
- [2] 이정연 경기도교육연구원, 연구보고서 "코로나19와 교육: 학교구성원의 생활과 인식을 중심으로", P. 126, 2020.08.31
- [3] 이정연 경기도교육연구원, 연구보고서 "코로나19와 교육: 학교구성원의 생활과 인식을 중심으로", PP. 129-132, 2020.08.31
- [4] 장채원, 신문 기사 "3D 프린터를 무료로 사용하는 방법", 대한민국 정책 브리핑, 2019.07.08
- [5] Musio, <https://www.themusio.com/home>
- [6] Google Dataset, <https://github.com/googlecreativelab/quickdraw-dataset>
- [7] G. Bradski, The OpenCV Library, Dr. Dobb's Journal of Software Tools, Priority 4, 2000.
- [8] Amazon Rekognition, <https://aws.amazon.com/ko/rekognition/>
- [9] 오재호 경기연구원, "코로나19가 앞당긴 미래, 교육하는 시대에서 학습하는 시대로", 이슈&진단 제 421호, PP. 1-25, 2020.06