

시각장애인을 위한 CNN 기반의 점자 변환 및 음성 출력 장치 설계

박승빈¹, 김봉현^{2*}

¹서원대학교 정보통신공학과 석사과정, ²서원대학교 컴퓨터공학과 교수

Design of CNN-based Braille Conversion and Voice Output Device for the Blind

Seung-Bin Park¹, Bong-Hyun Kim^{2*}

¹Master's Student, Department of Information Communication Engineering, Seowon University

²Professor, Department of Computer Engineering, Seowon University

요약 시대가 발전함에 따라 정보가 다양해지고 이를 얻는 방법도 다양해진다. 살아가면서 얻는 정보의 양 중 약 80%는 시각적 감각으로 습득한다. 하지만 시각장애인들은 시각 자료를 해석하는 능력이 제한된다. 그래서 점자라는 시각장애인용 문자가 등장했다. 그러나 시각장애인들의 점자 해독률은 5%에 불과하며 시간에 지남에 따라 다양한 형태의 플랫폼이나 자료를 원하는 시각장애인들의 요구가 늘어나면서 시각장애인들을 위한 개발 및 물품 제작이 이루어지고 있다. 물품 제작의 예로는 점자 도서를 들 수 있는데 이 점자 도서는 장점보단 단점이 많아 보이고 비장애인과 다르게 아직도 정보 접근에 대해서는 많이 어려운 것이 사실이다. 본 논문에서는 시각장애인이 정보를 기존의 방법보다 쉽게 얻을 수 있도록 CNN 기반 점자 변환 및 음성 출력 장치를 설계하였다. 이 장치는 점자로 되어 있지 않고 점자로 제작이 되지 않은 책, 텍스트 이미지나 손글씨 이미지 등을 카메라 인식을 통해 점자로 변환할 수 있도록 하고, 점자로 변환 후 시각장애인들의 요구에 따라서 음성으로 변환해 출력할 수 있는 기능을 설계해 시각장애인들이 정보를 얻을 수 있게 도와주어 삶의 질을 높이고자 한다.

주제어 : 시각장애인, CNN(Convolution Neural Network), 점자, 음성 인식, 사물인터넷

Abstract As times develop, information becomes more diverse and methods of obtaining it become more diverse. About 80% of the amount of information gained in life is acquired through the visual sense. However, visually impaired people have limited ability to interpret visual materials. That's why Braille, a text for the blind, appeared. However, the Braille decoding rate of the blind is only 5%, and as the demand of the blind who want various forms of platforms or materials increases over time, development and product production for the blind are taking place. An example of product production is braille books, which seem to have more disadvantages than advantages, and unlike non-disabled people, it is true that access to information is still very difficult. In this paper, we designed a CNN-based Braille conversion and voice output device to make it easier for visually impaired people to obtain information than conventional methods. The device aims to improve the quality of life by allowing books, text images, or handwritten images that are not made in Braille to be converted into Braille through camera recognition, and designing a function that can be converted into voice according to the needs of the blind.

Key Words : Blind, Convolution Neural Network (CNN), Braille, Speech Recognition, IoT

*교신저자 : 김봉현(bhkim@seowon.ac.kr)

접수일 2023년 4월 5일 수정일 2023년 5월 27일 심사완료일 2023년 5월 30일

1. 서론

정보화 사회에서 정보의 가치는 정말 중요하다. 정보 기술이 점점 발전하면서 일반적인 사람들뿐만 아니라 시각과 청각장애 등 여러 가지의 장애 유형의 사람들도 시각, 청각 그리고 촉각 등 오감과 느낌을 더하여 정보를 습득하고 판단할 수 있게 된다[1]. 그 중 살아가면서 얻는 정보의 양 중 약 80%는 시각적 감각으로 습득한다. 반대로 이 수치는 시각장애인들에게 있어 아주 치명적인 수치이다. 국가통계포털에서의 2021년 12월 말 전국 장애 유형 별 등록 장애인 수 현황에서 1위는 지체, 2위는 청각 마지막으로 시각장애인은 251,620명으로 전체 장애 등록인 수 중 3위로 높은 수치를 보인다. 시각장애인은 시각 자료를 해석하는 능력이 제한되어 있으며 점점 발전하는 정보사회에서 정보를 습득하고 활용하는 데 있어 다양한 방법이 있어도 시각장애인들의 접근이 어려운 것은 사실이며 이는 정보격차로 이어질 수 있다. 그래서 물리적 관점에서 데이터를 해석하는 것이 어려운 시각장애인을 위해서 점자가 등장했다. 점자란 종이, 지면, 물건 등 표면 위에 도드라진 점을 손가락의 촉각을 통해 읽을 수 있도록 하는 시각장애이용 문자이다. 이것을 통해 교육, 독서, 사회 참여 등 생활하는 데 있어 정보를 얻을 수 있다[2]. 말하기, 듣기 등 소통을 할 때 어려움이 없을 수 있지만 어떠한 문자를 읽고 쓰는 행위가 시각적인 요인 때문에 제한되어 점자를 사용한 언어활동을 한다. 따라서 시각장애인들에게 점자란 일상생활은 물론 삶에 있어서 필수적인 요소로 보인다[3]. 하지만 시각장애인들의 점자 해독률은 안타깝게도 약 5%에 불과하며 이는 후천적 장애가 많아 점자를 해석하지 못하는 경우가 많기 때문이다. 그래서 시간이 지남에 따라 다양한 형태의 플랫폼 품이나 자료를 원하는 시각장애인들의 요구가 늘어나면서 시각장애인들을 위한 물품 제작과 시각 감각을 대체한 촉각 그리고 청각 등을 활용하여 정보를 습득할 수 있는 개발이 진행되고 있으며 물품 제작의 예로 점자로 책을 출간하는 점자 도서를 들 수 있다[4]. 하지만 한국장애인단체총연맹에서는 점자 도서의 비율이 2021년 기준 출간한 도서의 0.2%라는 아주 작은 숫자에 불과하며 일반 도서의 5배 이상의 비싼 가격으로 해당 물품 사용에 있어 정보 접근에 다가갈 순 있어도 경제적 단점이 더 두드러진다[5]. 이에 본 논문에서는 점자 해독률과 점자 도서의 분포도를 검토 후 기존 점자를 해석하는데 필요한 촉각에 청각을 더한 장치를 개발하여 시각적인 자료 해석에 도움을 주어 시각장애인의 삶의 질을 향상에 목적이 있다[6].

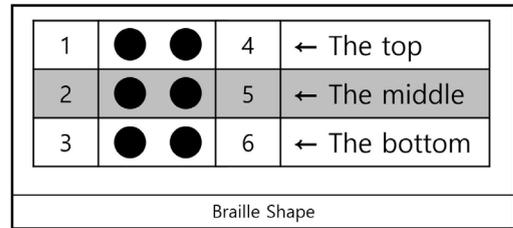
이 있다[6].

따라서 본 논문에서는 시각장애인의 정보력을 높이고 자 시각적 자료의 해석을 도와주기 위해 시각적 자료를 카메라로 인식하여 결과로 나온 글자를 점자로 변환을 하고 음성으로 출력할 수 있는 장치를 설계하고자 한다.

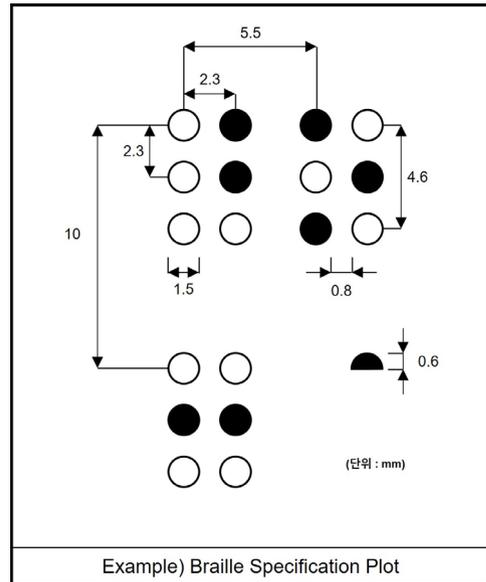
2. 관련이론

2.1 점자규정 및 규격

점자는 가로로 2점, 세로로 3점으로 총 6개의 점으로 구성되어 있으며 64개의 점형을 만들 수 있고 각각의 점형에는 의미가 부여된 문자이다.



[Fig. 1] A Braille Shape



[Fig. 2] Braille Specification Plot

2021년 6월 9일에 시행된 「점자법」 제3조에서는 ‘점자’란 시각장애인이 촉각을 활용하여 스스로 읽고

쓸 수 있도록 튀어나온 것을 일정한 방식으로 조합한 표기 문자를 말하며, 도형 및 그림 등을 촉각으로 인지할 수 있도록 제작된 촉각 자료를 포함한다고 명시되어 있다.

2020년 9월 10일, 문화체육관광부(이하 문체부)에서는 「한국 점자 규정」 2017년-15호를 개정·고시했다. 이번 개정에서는 점 높이를 0.6-0.9mm, 정의 지름을 1.5-1.6mm, 점 간격 2.3-2.5mm, 줄 간 간격 10.0mm 이상(최뎃값 정하지 않음), 자간 거리는 종이와 스티커의 경우 5.5-6.9mm, 피브이시(pvc)의 경우 5.5-7.3mm, 알루미늄과 스테인리스는 5.5-7.6mm이며 마지막으로 “기타 재질은 위의 재질의 규격을 필요한 수정을 가하여 적용” 등과 같은 물리적 규격을 제정해 점자 사용 기준을 신설했으며 위 Fig. 2와 같다.

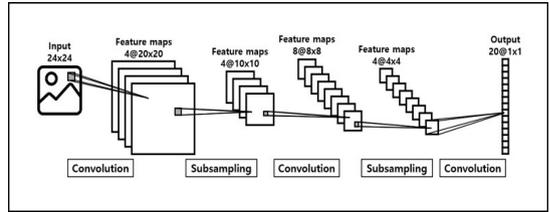
2.2 아두이노

아두이노(Arduino)는 저작권이 없는 오픈 소스를 기반으로 한 8비트 마이크로컨트롤러 단일 보드로 보드와 개발 도구 및 환경으로 소프트웨어와 하드웨어를 결합한 오픈소스 전자 플랫폼이다. 아두이노는 여러 가지의 센서나 스위치로 값을 받아 다른 센서나 스위치, 외부 장치를 통제할 수 있고 서로 간 상호작용하는 장치를 만들어 활용할 수 있다.

아두이노의 통합개발환경(IDE)에서 C-언어를 기반인 멀티 플랫폼으로 개발이 되는데 기존에는 C++ 언어를 사용하여 다양한 동작을 할 수 있도록 되어있지만 잘 갖춰진 라이브러리로 기초적인 C-언어 지식이 있어도 개발을 할 수 있어 누구나 쉽게 접근할 수 있는 장점이 있다. 개인용 컴퓨터와 아두이노 USB를 연결 후 마우스 클릭 한 번으로 컴파일 및 업로드가 가능한 컴파일러를 쉽게 할 수 있는 기능과 다양한 에디터를 통한 기능을 포함하고 있으며 www.arduino.cc에서 프로그램에 대한 도움을 쉽게 무료로 받을 수 있고 부품들이 저렴하다는 장점이 있다[7-9].

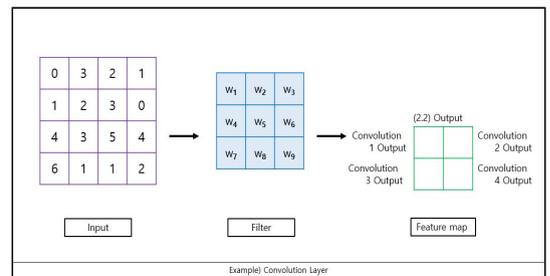
2.3 CNN

딥러닝 기법 중 이미지 인식 성능이 뛰어나고 주로 이미지나 음성인식 등에 사용되는 CNN은 합성곱 계층(Convolution Layer)과 서브 샘플링(Subsampling)이라고도 불리는 풀링 계층(Pooling Layer)으로 총 2개의 계층이 있다[10, 11].

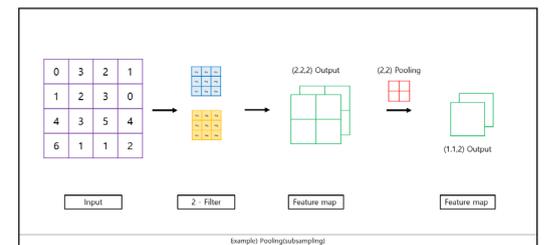


[Fig. 3] Typical CNN Structure

일반적인 CNN 구조는 이미지/영상이 입력되면 필터(Filter)를 통해 특징 맵(Feature map)을 만들고 특징을 추출하는 합성곱 계층과 풀링 계층(서브샘플링)을 쌓은 형태로 구성된 특징을 추출하는 영역(Feature Extraction)과 클래스를 분류하는 부분으로 나누어진다. 특징 추출 영역에서 합성곱 계층은 입력된 이미지나 영상 등 데이터에 필터 연산을 적용하여 특징을 추출한다. 추출된 특징들마다 필요에 따라서 서브 샘플링을 수행하는데 특징 특징들을 강조하는 효과를 얻을 수 있으며 풀링이라고도 한다. 풀링에는 최대 풀링(max pooling)과 평균 풀링(average pooling)이 있다. 예를 들어 (5,5)의 이미지가 입력되면 (3,3)의 필터를 적용할 경우 (3,3)의 산출(Output)을 내게 된다[8,9]. 또한 컨볼루션 커널(Convolution Kernel)의 계수를 설정하여 다양한 특징들을 추출할 수 있다[12].



[Fig. 4] Convolution Layer

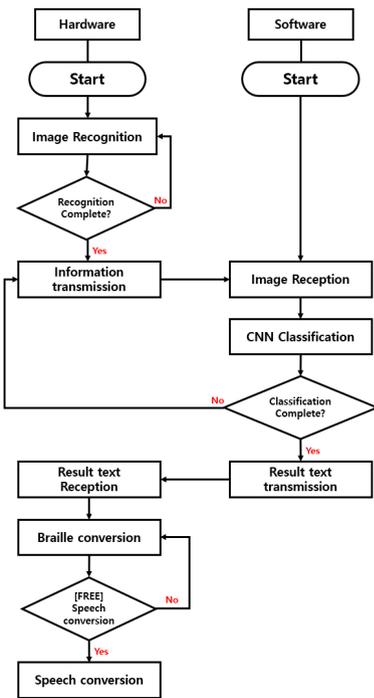


[Fig. 5] Pooling(subsampling)

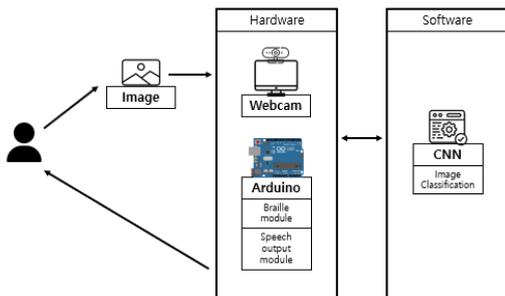
3. 점자 변환 및 음성 출력 장치 설계

3.1 시스템 설계도

본 논문에서 점자 변환 및 음성 출력 장치의 시스템 설계도는 다음 Fig 7과 같다. 사용자가 텍스트나 손글씨가 써진 종이를 웹캠과 아두이노 센서들로 이루어진 하드웨어에 전달하고 인식한 정보를 시리얼 통신으로 연결된 소프트웨어의 CNN 모델을 통해서 분류된 결과물을 하드웨어에 전달 후 결과를 점자로 변환하고 사용자가 자유롭게 음성으로 변환할 수 있도록 설계하였으며 순서도와 설계도는 다음 Fig 6, Fig 7로 나타내었다[13,14].



[Fig. 6] Device Flowchart



[Fig. 7] System Diagram

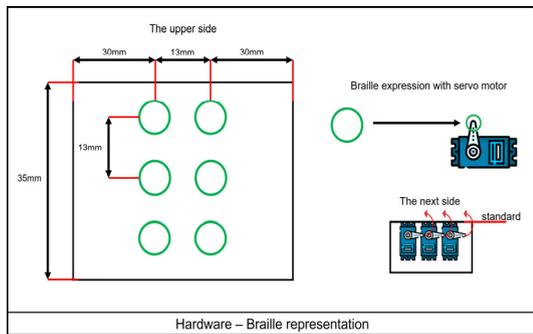
3.2 하드웨어 구성 및 설계

하드웨어를 구성하는 대표적인 부품들로는 아두이노 센서 제어에 위한 아두이노 우노 R3 보드, CNN 모델이 텍스트를 분류 결과를 점자를 표현하기 위한 서보 모터와 음성으로 출력하기 위한 음성 출력 모듈, 텍스트가 입력되었거나 써진 종이를 인식하기 위한 웹캠, 다수의 서보 모터를 제어하기 위한 어댑터와 이 어댑터를 연결하기 위한 여성 어댑터가 있다. 아래 Table 1은 하드웨어를 구성하는 부품들을 나타낸 표이다[15].

<Table 1> Hardware Parts

Hardware Parts		Purpose of Use
	Arduino Uno R3	sensor control
	Servo motor	Braille representation
	Voice output module	output Braille as voice
	Web cam	image recognition
	Adapter	current supply for multiple sub-motor control
	Adapter-female	adapter connection

본 논문에서는 점자를 아두이노 센서로 표현하기 위해서 소형 규격(22.2 X 11.8 X 31mm)의 서보 모터를 사용하였다. 서보 모터의 날개를 이용하여 점자의 불룩한 형태는 날개가 기준면 위로 향하여 나타났고 반대로 평평한 부분은 날개가 기준면 아래로 향하여 나타나도록 설계했다. 개정된 「한국 점자 규정」 2017년-15호에서 “기타 재질은 위의 재질의 규격을 필요한 수정을 가하여 적용”을 따르고 기본적인 점자 구성에는 가로로 2, 세로로 3점 총 6점으로 구성되어 있으므로 6개의 서보 모터를 사용하여 점자를 표현했다[16,17].



[Fig. 8] Hardware, Braille, representation

3.3 소프트웨어 설계

한글의 총 만들 수 있는 글자 수는 11,172개지만 실용성 있는 글자 수는 2,350개이다. 따라서 본 논문에서는 텍스트 종이를 인식하기 위해서 2,350개의 글자 수와 도서용 서체를 적용하여 여러 이미지 데이터와 손글씨는 캐글(Kaggle)에서 제공하는 숫자 0~9까지로 구성된 65,000개의 손글씨 데이터 셋(55,000개의 훈련 데이터 셋, 10,000개의 테스트 데이터 셋)으로 이루어진 이미지를 이용하여 CNN(Convolution Neural Network)으로 분류 후 결과를 하드웨어에 전달할 수 있도록 설계하였다[17,18].

4. 결론

살아가면서 오감과 느낌을 더해 정보를 습득하고 판단한다. 그 정보의 약 80%는 시각을 통하고 얻게 된다. 시각장애인은 시각 자료를 해석하는 능력이 제한되어 있어 점자라는 시각장애이용 문자가 등장했다. 하지만 점자 해독률은 5%밖에 되지 않고 시대가 변화함에 따라 다양한 형태의 정보를 얻을 수 있는 플랫폼이나 자료의 요구가 높아지면서 물품 제작과 기존 점자를 해석하기 위해 이용하는 촉각에 다른 감각을 더하거나 대체한 개발이 진행되고 있다.

본 논문에서는 위를 기반으로 시각장애인을 위한 CNN 기반 점자 변환 및 음성 인식 장치를 설계하였다. 시각장애인이 정보를 기존의 방법보다 쉽게 얻을 수 있도록 점자로 제작이 되어 있지 않은 텍스트 이미지, 손글씨를 점자로 변환하고 음성으로 들을 수 있는 기능을 사용해서 정보를 얻을 수 있도록 하였다. 현재 설계 시점에서는 점자 변환 및 음성인식을 수행하는 단계이지만, 향후 음성을 점자로 변환하고 점자 교육을 아두이노 센서

들로 할 수 있는 시스템으로 확장하여 점자 해독률을 높이면 점점 발전하는 정보화시대에 시각장애인들이 정보를 얻을 수 있도록 도와주어 삶의 질을 향상할 수 있게 하는 데에 목적이 있다.

REFERENCES

- [1] S. A. Hwang and S. M. Lee, "A Study on the Utilization of Alternative Materials for the Visually Impaired Using Reading Aids Devices," Journal of the Korean BIBLIA Society for library and Information Science, Vol.13, No.3, pp.191-211, 2020.
- [2] H. J. Kim, J. H. Moon, M. U. Song, S. M. Lee and K. S. Kong, "OnDot: Braille Training System for the Blind," The Institute of Internet, Broadcasting and Communication, Vol.20, No.6, pp.41-50, 2020.
- [3] Y. I. Kim and H. W. Moon, "An Examination of the Current Status and Needs of Braille Literacy Education for Individuals with Adventitiously Visual Impairments," The Korean Journal of Visual Impairment, Vol.33, No.1, pp.1-29, 2017.
- [4] M. J. Yoon, J. H. Kang and K. H. Yu, "A Study of Walking Guide for the Blind by Tactile Display," Journal of Institute of Control, Robotics and Systems, Vol.13, No.8, pp.783-789, 2007.
- [5] H. J. Kim, Y. C. Kim, C. J. Park, S. J. Oh and B. J. Lee, "Auto Braille Translator using Matlab," The Journal of the Korea institute of electronic communication sciences, Vol.12, No.4, pp.691-700, 2017.
- [6] J. Y. Lee and W. S. Cho, "A Study of Way guide System Design for a Visual Handicapped in School," Journal of the Korea Society of Computer and Information, Vol.15, No.4, pp.75-82, 2010.
- [7] S. H. Seo and S. W. Jang, "Design and Implementation of a smart shoes module based on Arduino," Journal of the Korea Institute of Information and Communication Engineering, Vol.19, No.11, pp.2697-2702, 2015.
- [8] H. S. Lee and J. C. Oh, "Design and Implementation of a Small Server Room Environment Monitoring System by Using the Arduino," The Journal of the Korea institute of electronic communication sciences, Vol.12, No.2, pp.385-390, 2017.
- [9] H. Y. Eom and K. H. Lee, "Design of Embodiment-based Programming Education using Arduino for Middle School Students," The Journal of the Convergence on Culture Technology, Vol.6, No.1, pp.471-476, 2020.
- [10] Y. G. Kim and E. Y. Cha, "Streamlined GoogLeNet Algorithm Based on CNN for Korean Character Recognition," Journal of the Korea Institute of Information and Communication Engineering, Vol.20, No.9, pp.1657-1665, 2016.

- [11] J. H. Park, J. A. Kim, H. J. Kim and D. H. Kim, "CNN based modeling and classification for variety of apples," Journal of D-Culture Archives, Vol.4, No.1, pp.63-70, 2021.
- [12] T. Q. Hieu, S. H. Yeon and J. M. Kim, "Korean License Plate Recognition Using CNN," Journal of IKEEE, Vol.23, No.4, pp.1337-1342, 2019.
- [13] S. H. Choi, J. H. Kim, J. D. Oh and K. S. Kong, "A Smart Closet Using Deep Learning and Image Recognition for the Blind," The Journal of the Institute of Internet, Broadcasting and Communication, Vol.20, No.6, pp.51-58, 2020.
- [14] H. G. Kim and Y. G. Choi, "Study on the Development of Working Safety Device for Visually Impaired Person," The Journal of Korea Institute of Information, Electronics, and Communication Technology, Vol.9, No.4, pp.366-372, 2016.
- [15] G. Y. Heo, "Implementation of an Arduino Compatible Modular Kit for Educational Purpose," Journal of the Korea Institute of Information and Communication Engineering, Vol.23, No.5, pp.547-554, 2019.
- [16] C. G. Kim and J. M. Seo, "An Design and Implementation of Navigation System for Visually Impaired Persons Based on Smart Mobile Devices," The Journal of the Korea Contents Association, Vol.15, No.1, pp.24-30, 2015.
- [17] S. H. Choi and S. H. Jung, "Performance Improvement of Fake Discrimination using Time Information in CNN-based Signature Recognition," Journal of Digital Contents Society, Vol.19, No.1, pp.205-212, 2018.
- [18] N. K. Lee, J. Y. Kim and J. H. Shim, "Empirical Study on Analyzing Training Data for CNN-based Product Classification Deep Learning Model," The Journal of Society for e-Business Studies, Vol.26, No.1, pp.107-126, 2021.

김 봉 현(Kim, Bong Hyun)

[정회원]



- 2009년 2월 : 한밭대학교 컴퓨터 공학과 (공학박사)
- 2012년 3월 ~ 2015년 2월 : 경남대학교 컴퓨터공학과 교수
- 2020년 3월 ~ 현재 : 서원대학교 컴퓨터공학과 교수

<관심분야>

Big data, IoT, AI Service, ICT convergence

박 승 빈(Park, Seung Bin)

[준회원]



- 2022년 2월 : 서원대학교 컴퓨터 공학과(공학사)
- 2022년 3월 ~ 현재 : 서원대학교 정보통신공학과 석사과정

<관심분야>

딥러닝 알고리즘, 데이터사이언스, 빅데이터