

# SLAM을 활용한 병원 자율주행 서비스 로봇

Autonomous Service Robot for Hospitals Using SLAM Technology

김서연, 남현수, 임승원





# 목차

## 1. 서론

### 1.1 기획 배경

## 2. 본론

### 2.1 주요 기능 및 흐름도

### 2.2 프로젝트 특/장점

### 2.3 다학제적 접근

### 2.4 자율 주행

### 2.5 얼굴 랜드마크 인식

### 2.6 음성 처리 기술

### 2.7 GUI 인터페이스

## 3. 사업화 방향

## 4. 미래 신산업 부합도

## 5. 결론



# 1. 서론

## 1.1 기획 배경

- 의료 토털 서비스 로봇 개발의 필요성

### 의료진 측면

- ① 의료진은 인력 부족으로 인하여 의료업무에 더해 다양한 업무를 맡고 있음.
- ② 반복 업무 및 서비스의 경우 로봇으로 대체·자동화 가능함.

→ AI 로봇을 도입하여 의료진의 업무 부담 ↓, 병원 운영 효율화

### 환자 측면

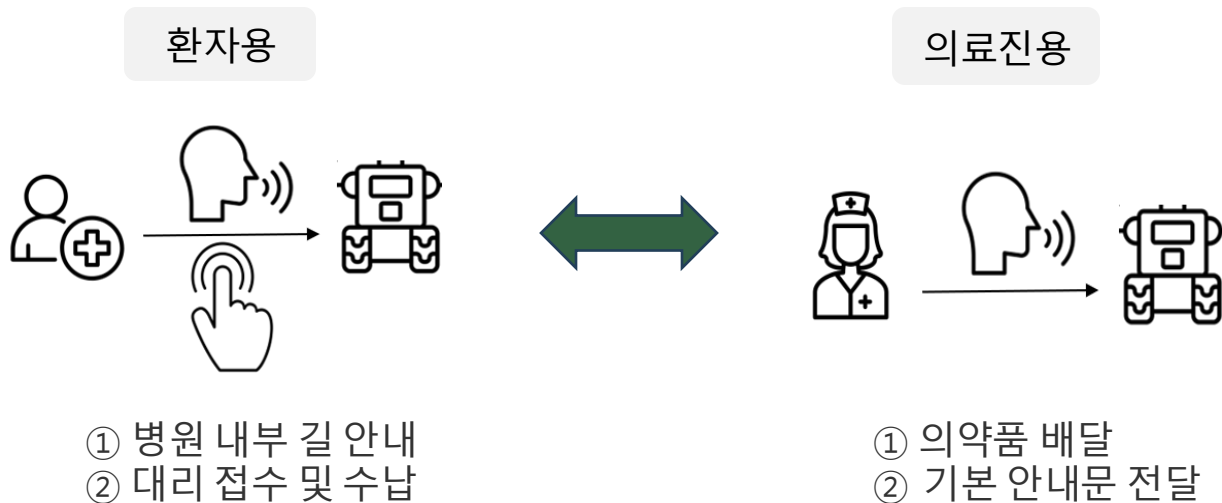
- ① 병원은 유동 인구가 많은 환경이며 환자들은 다양한 서비스를 필요로 함.
- ② 고령자를 포함한 다양한 유형의 환자가 동등하고 편리하게 의료 서비스를 이용할 수 있어야 함.

→ 의료 접근성 높여 모든 환자에게 친화적 의료 서비스 제공

## 2. 본론

### 2.1 주요 기능 및 흐름도

- 환자용 · 의료진용 서비스 제공 및 인터페이스 상호 전환
- 인공지능 기반 기술을 활용한 다양한 기능 탑재

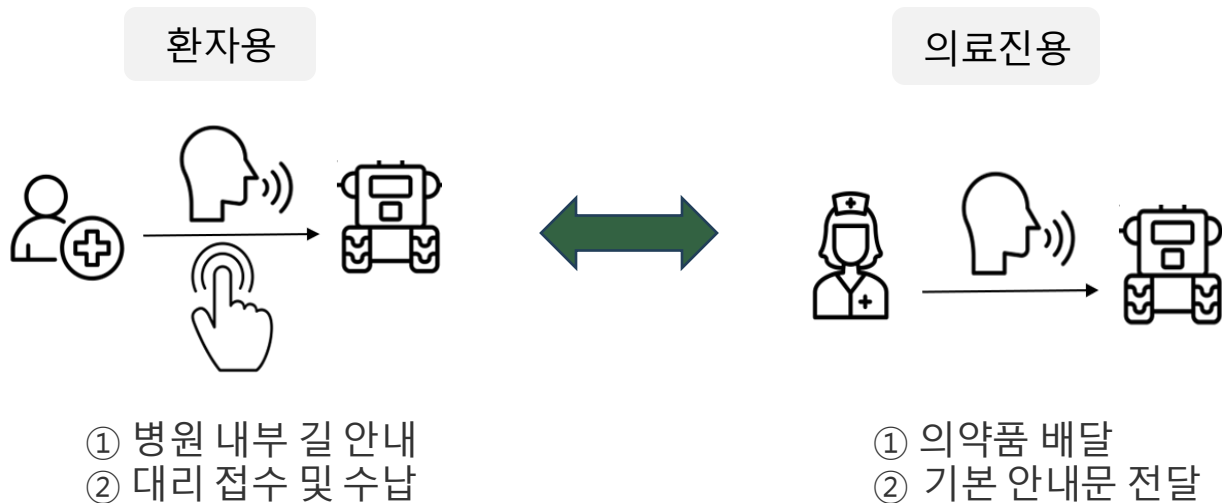


적용 기술: Face detection, STT(Speech-to-Text), TTS(Text-to-Speech), 자율 주행

## 2. 본론

### 주요 기능 및 흐름도

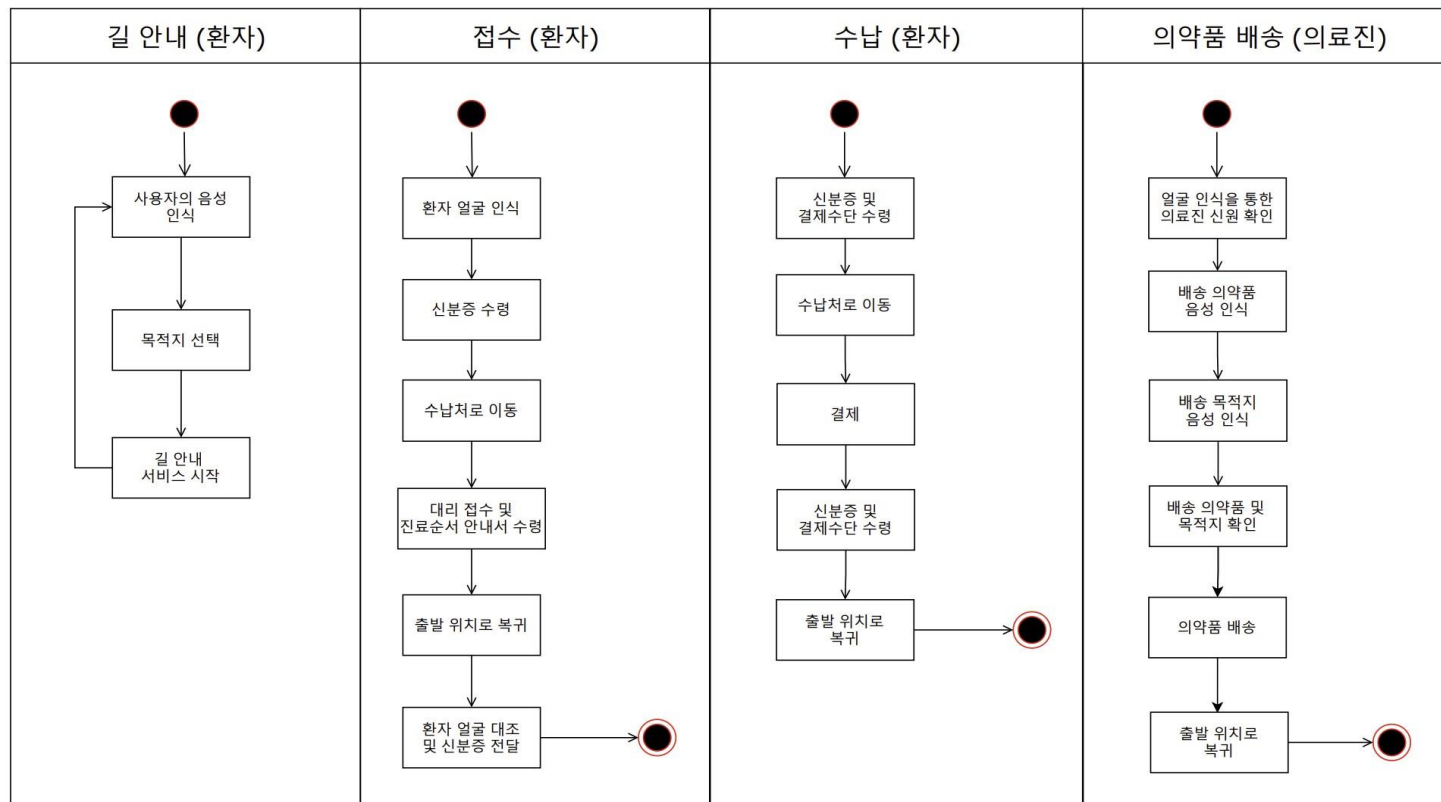
- 환자용 · 의료진용 서비스 제공 및 인터페이스 상호 전환
- 인공지능 기반 기술을 활용한 다양한 기능 탑재



적용 기술: Face detection, STT(Speech-to-Text), TTS(Text-to-Speech), 자율 주행

# 2. 본론

## 2.1 주요 기능 및 흐름도



# 2. 본론

## 2.2 프로젝트 특/장점

### 1) 토탈 서비스 로봇

- 단순 업무 보조, 안내 역할을 수행하는 기존 병원 로봇에 비해 다양한 서비스를 제공

### 2) 대상에 맞춘 개별화된 서비스

- 환자 및 의료진 모드를 동시에 제공하는 HCI 인터페이스로 다양한 기능 수행
- 하나의 소프트웨어에서 환자와 의료진을 구분하여 사용자 맞춤형 서비스를 제공

### 3) 얼굴 인식과 잠금장치가 결합된 서비스

- 얼굴 인식과 잠금장치가 결합된 서비스로 각 기능의 보안성 확보

# 2. 본론

## 2.2 프로젝트 특/장점

### 4) 사용자와 상호작용

- 음성 인식 기술을 통해 사용자와 로봇 간 상호작용

### 5) 자율주행

- SLAM 기반으로 병원 내 환경을 탐색하고 실시간으로 맵을 생성하여, 경로 탐색 및 최적 경로 계획으로 장애물을 회피하며 목적지까지 이동 가능



## 2. 본론

### 2.3 다학제적 접근

1



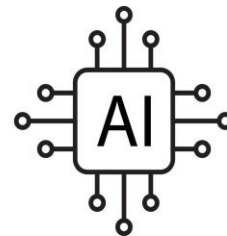
음성 처리 기술

=



자연어 처리

+



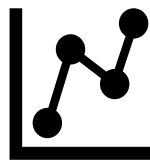
인공지능

2



얼굴 랜드마크 인식

=



통계학

+



컴퓨터 비전

## 2. 본론

### 2.3 다학제적 접근

3



GUI

=



HCI

+



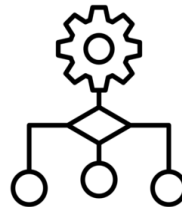
심리학

4



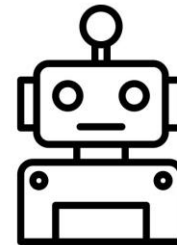
자율주행

=



임베디드 SW

+

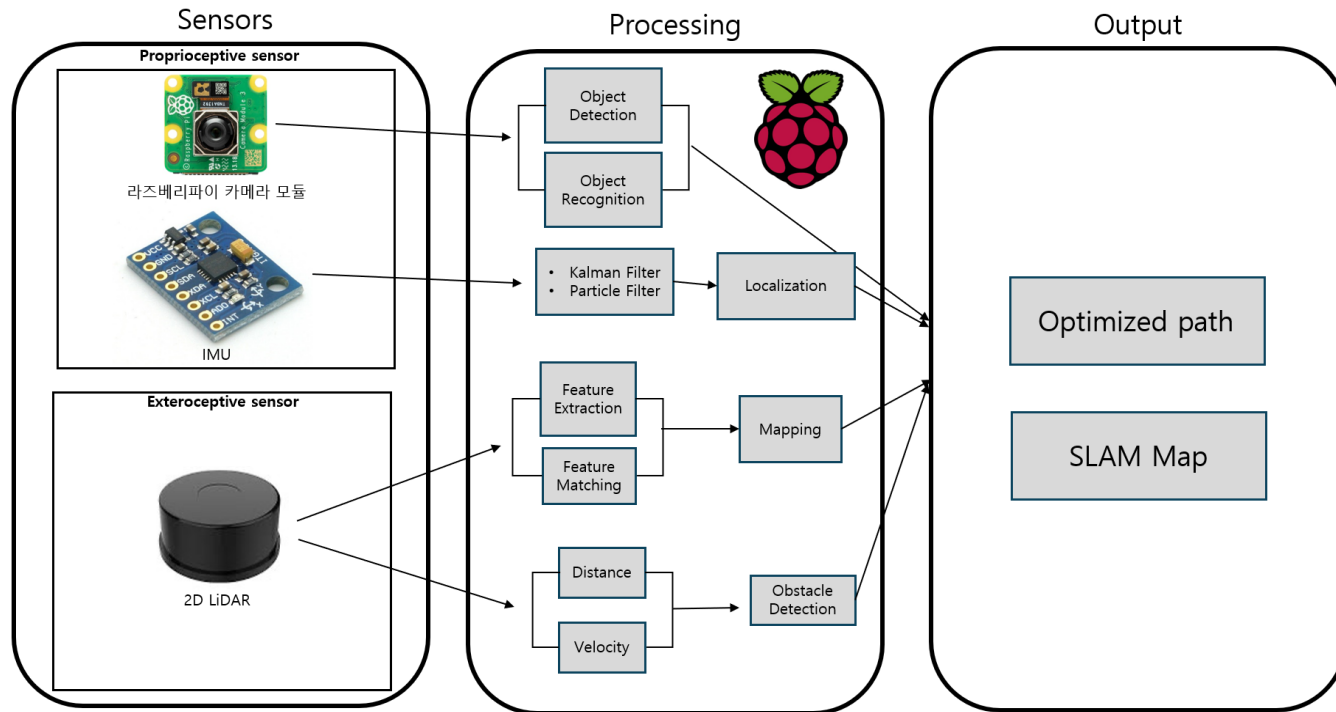


로봇 공학

## 2. 본론

### 2.4 자율 주행

- SLAM 맵 구성 시스템 구조도



## 2. 본론

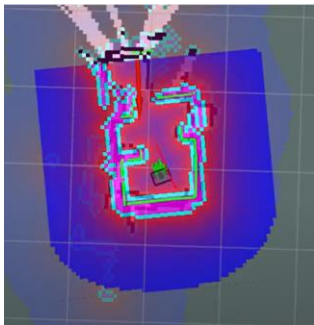
### 2.4 자율 주행



SLAM 지도 작성

#### SLAM(Simultaneous Localization and Mapping)

- 로봇이 실시간으로 위치를 추정하는 동시에 그 환경의 지도를 작성하는 기술
- 라이다(Lidar) 센서: 레이저를 사용한 고정밀 거리 측정으로 2D 지도 생성
- IMU(관성 측정 장치): 로봇의 움직임과 회전을 측정하여 정밀한 위치 추정



Navigation

#### DWA(Dynamic Window Approach) & Localization

- 장애물 회피와 최단 경로 이동을 고려한 최적의 경로 계획
- odometry 정보와 출발 위치를 바탕으로 맵에 Localization하여 현재 위치 추정
- 최적 경로로 주행하기 위한 속도와 각속도를 실시간으로 계산

## 2. 본론

### 2.5 얼굴 랜드마크 인식

- 목적: 신원 확인
- 적용 기능: 의료진 대상 기능, 환자 대상 '접수' 기능
- Mediapipe: 얼굴 특징점 468개를 추출
- FAISS: L2 거리로 특징점 벡터 간 유사도 측정
$$d(p, q)^2 = (q_1 - p_1)^2 + (q_2 - p_2)^2$$
L2 거리 식
- 유사도 < 임계값 → 신원 일치



병원 내 건강보험 본인  
확인 의무화 안내



얼굴 인식 GUI  
하며

## 2. 본론

### 2.5 얼굴 랜드마크 인식

- 임계값에 따른 얼굴 인식 성능 차이를 비교 **실험**
  - 1) 본인의 얼굴 랜드마크를 추출
  - 2) 저장된 랜드마크와 본인의 얼굴을 200번씩 비교
  - 3) 다른 참가자의 랜드마크와 본인의 얼굴을 100번씩 비교 \*독립시행 가정

## 2. 본론

### 2.5 얼굴 랜드마크 인식

#### 성능 평가 지표

- **True Positive (TP):** positive로 정확하게 예측한 것  
⇒ 기준인 사람 결과가 신원일치가 나오는 경우
- **False Positive (FP):** 실제로 negative인 값을 positive로 잘못 예측한 것  
⇒ 기준이 아닌 사람 결과가 신원 일치가 나오는 경우
- **True Negative (TN):** negative로 정확하게 예측한 것  
⇒ 기준이 아닌 사람 결과가 신원 불일치가 나오는 경우
- **False Negative (FN):** 실제로 positive인 값을 negative라고 잘못 예측한 것  
⇒ 기준인 사람 결과가 신원 불일치가 나오는 경우

## 2. 본론

### 2.5 얼굴 랜드마크 인식

- 실험 결과

임계값 ↑ → 유사도 판단 기준 완화 임계값 ↑ → Accuracy, Precision, F1-Score ↓, Recall ↑

신원 확인의 목적 ⇒ 보안

∴ 최적의 임계값: 0.0001

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

임계값	0.0001	0.0002	0.00025
TP(True Positive)	172	183	195
FP(False Positive)	2	37	164
TN(True Negative)	198	163	36
FN(False Negative)	28	17	5
Accuracy	0.925	0.865	0.578
Precision	0.989	0.832	0.543
Recall	0.860	0.915	0.975
F1-Score	0.920	0.872	0.698

표 1. 임계값에 따른 얼굴 인식 성능평가 지표



## 2. 본론

### 2.6 음성 처리 기술

#### STT (Speech to Text, 음성 인식)

- 사용자의 요구사항 또는 질문을 음성 데이터로 받아와 텍스트로 변환하는 기술
- speech\_recognition 모듈은 음성 인식을 위한 라이브러리로, Google Web Speech API를 사용하여 음성을 텍스트로 변환

```
# 음성 인식 함수
def recognize_speech(): 1개의 사용 위치
    recognizer = sr.Recognizer()
    with sr.AudioFile(WAVE_OUTPUT_FILENAME) as source:
        audio = recognizer.record(source)

    try:
        recognized_text = recognizer.recognize_google(audio, language="ko-KR") # 한국어 설정
        print(f"Recognized Text: {recognized_text}")
    except sr.UnknownValueError:
        print("Google Web Speech API could not understand the audio")
        recognized_text = ""
    except sr.RequestError as e:
        print(f"Could not request results from Google Web Speech API; {e}")
        recognized_text = ""

    return recognized_text
```



`recognizer.recognize_google(audio, language="ko-KR")`

Google Web Speech API에 음성 데이터를 전송해  
한국어로 변환된 텍스트를 반환

## 2. 본론

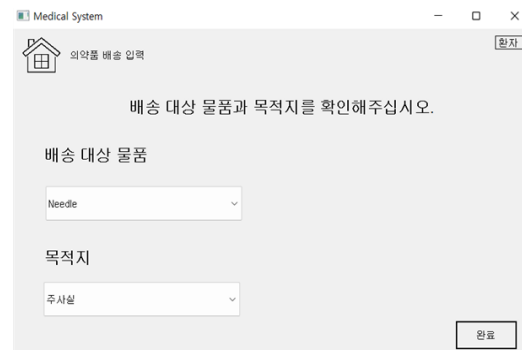
### 2.6 음성 처리 기술

#### STT (Speech to Text, 음성 인식)

- 적용 기능: 환자 대상 길안내, 의료진 대상 '의약품 배송'
- 사용자의 음성 데이터를 추출하여 길 안내를 제공받을 목적지, 배달할 의약품을 음성 데이터로 수집
- Google Speech API 및 후처리를 통하여 높은 정확도의 음성 인식을 구현
- 음성 인식 후, 인식된 물품과 목적지를 확인하는 GUI 페이지를 통해 재확인하여 정확도 향상



[그림1] 마이크로 음성 정보 수집



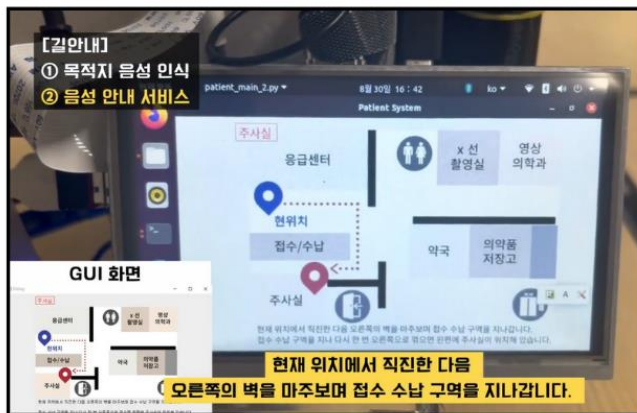
[그림2] 배송 물품과 목적지 확인

## 2. 본론

### 2.6 음성 처리 기술

#### TTS (Text to Speech, 음성 합성)

- 텍스트 데이터를 자연스러운 음성으로 변환하는 기술
- 로봇은 안내, 설명, 답변 등 다양한 텍스트 형태의 정보를 음성으로 제공
- Google Text-to-Speech API를 활용
- 사용자의 요청 사항에 대한 답변을 텍스트로 생성한 후 스피커를 통해 음성으로 안내



[그림3] 스피커로 길 안내 서비스 제공

## 2. 본론

### 2.6 음성 처리 기술

```
import pyaudio
import wave
import speech_recognition as sr
import pyttsx3

# 음성 녹음 설정
FORMAT = pyaudio.paInt16
CHANNELS = 1
RATE = 44100
CHUNK = 1024
WAVE_OUTPUT_FILENAME = "output.wav"

# 길찾기 대상 단어 설정
TARGET_WORDS = ["보호자 대기실", "화장실", "A 병동", "B 병동", "C 병동", "약국", "엘리베이터"]
RESPONSE_YES = "길 안내 기능을 제공드리겠습니다."
RESPONSE_NO = "길찾기 기능을 제공할 수 없습니다."
PROMPT = "목적지를 말씀해주세요."

# 음성 합성기 초기화
engine = pyttsx3.init()
```

pyttsx3: 로컬에서 동작하는 TTS 엔진.  
인터넷 연결 없이 텍스트를 음성으로  
변환

pyaudio: 음성 녹음 기능을 구현하는 데  
사용, 마이크로부터 실시간으로 음성을  
캡처해 output.wav 파일로 저장

```
# 메인 함수
def main(): 1개의 사용 위치
    # 사용자에게 목적지를 말해달라고 요청
    engine.say(PROMPT)
    engine.runAndWait()

    # 음성 녹음 및 인식
    record_audio()
    recognized_text = recognize_speech()

    # 인식된 텍스트 후처리
    recognized_text = preprocess_text(recognized_text)

    # 인식된 텍스트에 따라 응답
    if any(preprocess_text(word) in recognized_text for word in TARGET_WORDS):
        print(RESPONSE_YES)
        engine.say(RESPONSE_YES)
    else:
        print(RESPONSE_NO)
        engine.say(RESPONSE_NO)
    engine.runAndWait()

if __name__ == "__main__":
    main()
```

**engine.say(PROMPT)**

텍스트 형식으로 준비된 로봇의  
질문/답변/안내멘트 등을 음성으로 변환해 출력

# 2. 본론



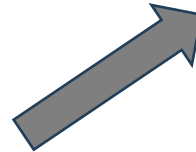
## 2.7 GUI 인터페이스



환자용 '병원 내 길찾기'와 '접수 및 수납' 을 위한 Home 화면

※ PyQt: Qt의 레이아웃에 Python의 코드를 연결하여 GUI 프로그램을 만들 수 있게 해주는 프레임워크

길찾기 선택시



사용자의 음성을 인식하여 목적지에 대한 최단 경로를 음성 및 텍스트로 안내함



## 2. 본론



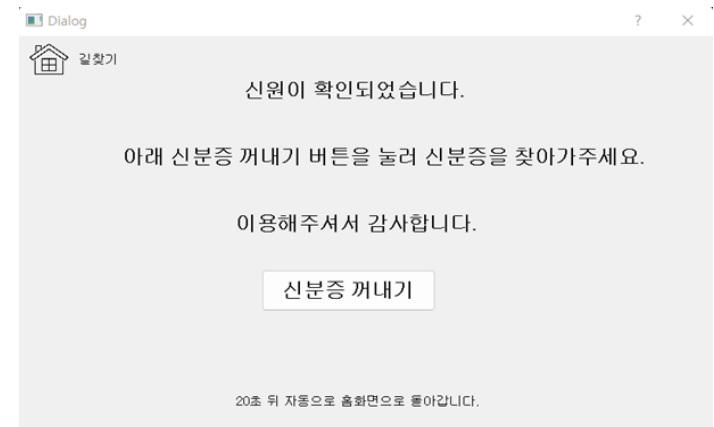
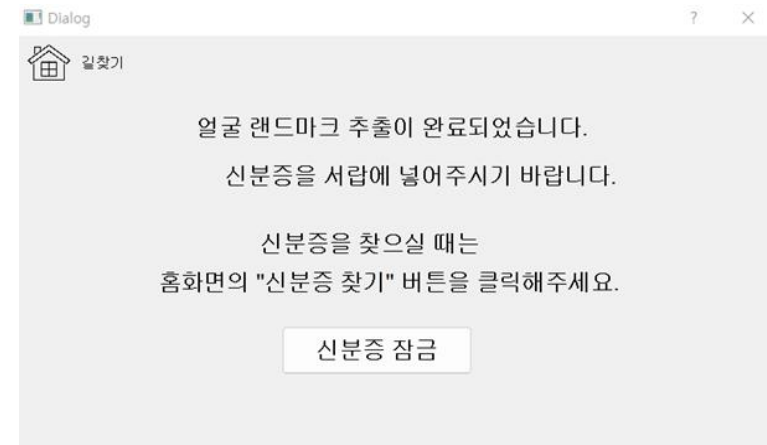
### 2.7 GUI 인터페이스

접수 선택 시, 얼굴인식을 통한  
신원 인식을 위해 카메라를 이용함.



신분증 찾기 선택 시, 얼굴 랜드마크  
추출을 위해 카메라를 이용함

랜드마크를 추출 완료 후의 화면



# 3. 사업화 방향

전국의 상급 종합병원 및 종합병원에 적용

환자 수가 많은 병원에 도입하여  
의료진의 업무 부담 경감



지역 경제 활성화에 이바지

로봇 지원 센터 설립 및  
지역 인재를 활용한 직무 창출



지역 의료 균형 발전에 기여  
고품질 의료 서비스 제공



충북대학교병원

글로벌 병원 시장 진출

국내 병원에서 성공적으로  
활용된 후 해외 진출



시장 확장 가능성

대학병원 뿐만 아니라  
요양병원, 재활병원 등  
다양한 의료 시설에 도입



# 4. 미래 신산업 부합도

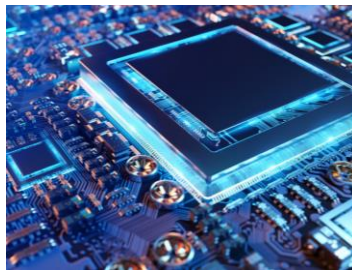
## 5대 신성장 산업



이차전지



전기차



반도체



디스플레이



바이오 헬스

## 5대 신성장 사업 관련 병원 토탈 서비스 로봇 기능

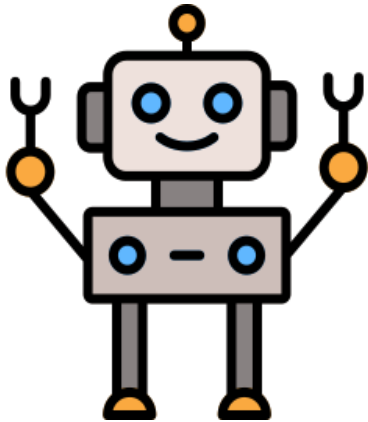
- 전기 에너지로 충전되어 이동
- 반도체를 이용하여 센서 데이터 처리 및 움직임 제어
- 로봇과 사용자 간 상호작용을 위한 디스플레이 탑재
- 환자와 의료진 모두의 편의를 증진하는 기술

⇒ 의료진의 작업을 효율적으로 지원, 환자 치료에 기여하여 의료 서비스 수준 제고

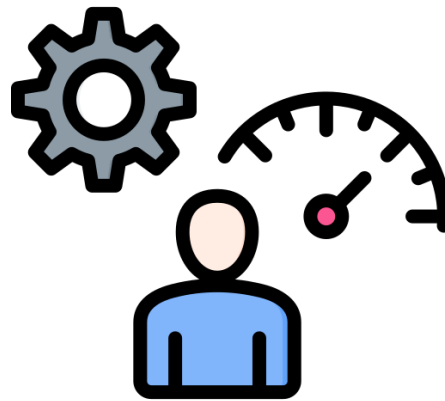


# 5. 결론

의료 토탈 서비스 로봇



의료진의 업무 부담 감소



병원 운영 효율화

