

컴퓨터 비전 기술을 활용한
시각장애인 금융 활동 지원 서비스
A Financial Assistance Service for the Visually
Impaired Using Computer Vision Technology



지도 교수 : 전상률

팀명 : 첫눈

202055557 성민기

202055556 성도범

202055559 송시우

목차

1. 과제 목표

- 과제 배경
- 과제 목표

2. 대상 문제 및 요구사항 분석

- 유사 시스템 분석
- 문제점 분석
- 시스템 필요성

3. 시스템 구성

- 시스템 구성도
- 개발 환경
- 주요 기능
- 구현 방식

4. 개발 일정 및 역할 분담

- 개발 일정
- 역할 분담

1. 과제 목표

과제 배경

시각장애인에 대한 전 세계적인 통계에 따르면, **2억 9,500만 명은 중등도 이상의 시각장애**를 겪고 있는 사람들로, 이들은 부분적인 시력 상실을 경험하거나 시력에 심각한 제한이 있습니다. 이들은 여전히 어느 정도 시각적 기능을 이용할 수 있지만, 일상생활에서 큰 불편을 겪고 있으며, 시각 보조 장치나 기술을 통해 활동을 이어가고 있습니다.

그 외에도, **4,330만 명은 중증 시각장애인**으로, 거의 또는 완전히 시력을 잃은 사람들입니다. 이들은 일상생활에서 시각을 거의 전혀 사용하지 못하며, 점자나 음성 안내 시스템 등을 통해 생활을 영위하고 있습니다. 중증 시각장애인들은 일상적인 활동에 큰 어려움을 겪고 있으며, 완전한 자립적인 생활을 위해 더 많은 지원이 필요합니다.

한국에서는 2021년 기준으로 약 **26만 명이** 등록된 시각장애인으로, 그 중 상당수가 후천적인 원인으로 시력을 상실한 사람들이 많습니다. 시각장애인들은 일상적인 생활에서 많은 제약을 겪고 있으며, 특히 **금융 활동**에 있어 큰 어려움을 겪고 있습니다. 기존의 금융 서비스는 대부분 **대면 서비스** 중심으로 제공되고 있으며, 이는 시각장애인에게 매우 불편한 상황을 초래합니다. 예를 들어, 시각장애인은 은행을 방문하거나, 보호자와 동행하여 서류를 처리하는 등의 불편함을 겪고 있습니다.

이러한 상황을 해결하기 위한 다양한 금융 서비스들이 제시되고 있지만, 대부분 점자 서류 제공, 음성 변환된 계약서 제공, 음성 안내 기능이 탑재된 ATM, 그리고 금융 상품 가입 시 보호자 동행 없이 혼자서도 금융 업무를 볼 수 있도록 지원하는 것에 그치고 있습니다. 이러한 대면 중심의 서비스는 시각장애인이 **온라인 금융 서비스**에서 제공되는 다양한 편리함을 누릴 수 없는 현실을 여전히 남겨두고 있습니다.

또한, **금융 문맹** 문제도 심각합니다. 많은 사람들이 금융 상품의 특성이나 조건을 정확히 이해하지 못하고 있으며, 이는 잘못된 금융 상품 선택이나 관리의 어려움을 초래하고 있습니다. 특히, 시각장애인들은 정보 접근성이 제한되어 있어 이러한 문제에 더욱 취약합니다.

더불어, **금융 문서의 과다한 양과 복잡한 관리**도 큰 문제로 지적되고 있습니다. 다양한 금융 상품과 서비스로 인해 수많은 문서가 발급되며, 이를 관리하는 데 어려움을 겪는 경우가 많습니다. 특히, 발급일, 납부일 등 중요한 일정 정보를 놓치게 되어 불이익을 받는 사례도 발생하고 있습니다.

과제 목표

본 과제는 시각장애인이 단순히 대면 서비스를 이용할 때뿐만 아니라 **온라인 금융 서비스**를 더욱 **편리하고 안전하게** 이용할 수 있도록 지원하는 시스템을 개발하고자 합니다. 음성 인식(STT), 텍스트 음성 변환(TTS), 그리고 광학 문자 인식(OCR) 기술을 활용하여, 시각장애인이 **자율적으로 금융 거래를 처리하고, 문서 관리 및 일정 알림을 받을 수 있는 시스템**을 제공할 예정입니다. 이 시스템은 시각장애인이 온라인 환경에서도 손쉽게 금융 서비스를 이용할 수 있는 기반을 마련해줄 것입니다.

2. 문제 및 요구사항 분석

유사 시스템 분석

구분	 KB국민은행	 신한은행	 NH농협은행	 토스	 '첫 눈' (본 프로젝트)
음성 안내	스크린리더 호환 웹/앱, ARS 음성인식 도입, 음성지원 OTP 발급	조력제도 통해 직원이 음성으로 주요 내용 안내, ATM 이어폰 음성안내 지원	해피뱅킹(화면 음성읽기) 제공, ATM 음성안내 기능 점진 도입, AI 음성인식 ATM 시범	앱 전반 VoiceOver 대응, 상담 연결 시 즉시 전화연결, ARS단계 화면 대체	AI 음성비서 내장 (음성명령으로 조회/송금), OCR로 문서내용 실시간 음성변환, 음성으로 양방향 소통 인터페이스 지원
점자 지원	점자블록 설치 (ATM/출입구), 점자 보안카드·카드·통장 제공, 약관 등 점자 번역	‘마음맞춤 키트’ 내 점자스티커·지폐 가이드, 점자 통장·보안카드 지원 (기본 제공)	점자 보안카드 (2011~), 점자 통장/명세서 요청 가능, 금융계약서 점자/텍스트 제공 추진	해당 없음 (지점·서류 없음), 카드번호 미각인 등 디자인으로 배려	점자보다 음성중심 접근, 필요시 연동 점자 키보드 입력 가능
앱/IT 접근성	WCAG 2.1 준수, 웹와치 인증, 모바일 앱 기본 제공	접근성 고려 UI (글자확대, VoiceOver 지원 등) 도입, 수어·자막 등 멀티모달 서비스	WCAG 2.0 인증, 웹/앱 챗봇 개선, 생체인증 ATM 개발 등 ICT 활용	모바일 접근성 우수앱 인증 , 동적 글자크기·다크모드 완비, 시각장애인 사용자 경험 풍부	AI 기술 총합 활용 , 멀티뱅크 데이터 통합, 사용자 맞춤형 UX 설계 (시각장애 특화)

비대면 거래 편의	일부 업무는 상담원 연계 필요, 음성 OTP로 비대면 인증 개선	조력제도로 대면 보조 지원, 완전 비대면 가입은 각 은행 앱 이용	접근성 높은 앱으로 계좌개설 등 비대면 가능하나, 문제시 직원 도움 필요	지점 없는 100% 비대면 , 앱 하나로 계좌개설·대출 등 완료, 상담도 24시간 가능	처음부터 끝까지 비대면 설계, 여러 은행 계좌를 한 앱에서 관리, 보조 인력 없이 독립적 금융생활
------------------	-------------------------------------	---	--	---	--

문제점 분석

대면 중심 접근성 의존

- 많은 시중은행은 여전히 시각장애인 고객을 위해 **점자 안내서나 직원 보조**에 의존하고 있습니다.
- 예를 들어, 신한은행의 ‘조력제도’는 혼자 방문한 시각장애인에게 친절하 서비스 제공하지만, **기본적으로는 창구 방문을 전제로 하며 비대면으로는 제공되지 않습니다.**
- KB국민은행, 농협은행 등도 대부분의 점자/음성 서비스는 **영업점에서 요청 시** 제공되는 방식입니다.

서비스 간 단절

- 각 은행은 자사 고객을 대상으로만 서비스를 운영하고 있어, **여러 은행에 계좌를 보유한 시각장애인이 일관된 접근성을 확보하기 어렵습니다.**
- 사용자 입장에서는 **은행별 앱, 웹사이트, 문서 포맷, 접근성 기능을 모두 따로 익혀야 하는 불편**이 발생합니다.

제한적인 디지털 접근성

- 앱이나 웹의 접근성은 점차 개선되고 있으나, 여전히 일부 메뉴는 **스크린 리더로 인식되지 않거나**, 음성으로 수행할 수 없는 절차(예: 본인 인증)가 존재합니다.

- 점자 OTP나 점자 보안카드는 인쇄물로 제공되며, 디지털 인증과는 **분리된 장치**로 사용되고 있습니다.

보조적 기능에 머무는 기술 활용

- 시중은행은 대부분 음성 안내나 점자 출력처럼 **보조적 수단**을 제공하는 수준입니다.
- **AI나 실시간 해석 기능은 거의 활용되지 않고**, 단순히 화면낭독 기능을 지원하거나 ATM에 음성 안내를 추가하는 방식에 그치고 있습니다.

비대면 서비스의 제한

- 토스를 제외하면 대부분의 은행은 시각장애인이 **혼자서 완전히 비대면으로 금융 업무를 처리하기 어렵습니다.**
- 특히 약관 확인, 문서 작성, 본인 인증 등은 **보조인의 도움을 전제로 한 설계**가 많아 **자립적 금융 생활에 장벽**이 존재합니다.

시스템 필요성

현재 시중은행들은 시각장애인을 위한 다양한 서비스를 제공하고 있지만, 여전히 대면 중심의 접근성에 의존하거나 보조적인 수준에 머무르고 있는 경우가 많습니다. 점자 안내서, 음성 OTP, 점자 보안카드 등은 고객이 직접 영업점을 방문하거나 요청해야만 제공되는 경우가 대부분이며, 각 은행의 서비스가 분절되어 있어 사용자는 금융기관마다 상이한 접근 방식과 절차를 반복적으로 학습해야 하는 불편을 겪고 있습니다. 또한, 모바일 앱이나 웹사이트 접근성 역시 일부 메뉴나 인증 절차에서 화면낭독 기능이 제대로 작동하지 않거나, 음성 명령을 통한 실질적인 조작이 어려운 경우도 존재합니다.

이처럼 시중 서비스가 비대면 자율 금융의 구현에는 아직 도달하지 못한 현실 속에서, ‘첫 눈’ 서비스는 기술적으로 한 발 더 나아간 해법을 제시합니다. ‘첫 눈’은 음성인식(STT), 음성안내(TTS), 문서 인식(OCR) 등 AI 기반 기술을 활용하여 시각장애인이 누구의 도움 없이도 금융 문서를 읽고 이해하며, 음성 명령만으로 주요 기능을 실행할 수 있도록 설계되었습니다. 단순한 보조 수단이 아닌, 처음부터 시각장애인의 금융생활 전반을 자립적으로 설계할 수 있도록 돕는 점에서 기존 서비스와의 본질적인 차이를 갖습니다.

따라서 ‘첫 눈’은 여전히 제한적인 시중은행의 접근성 서비스를 보완하며, 시각장애인이 진정으로 독립적인 금융활동을 할 수 있도록 지원하는 **필수적인 디지털 금융 접근성 플랫폼**을 제공하고자 합니다.

3. 시스템 구성

시스템 구성도(1)



- 음성 명령

- “이체”, “문서 스캔”, “일정 알려줘” 등

- OCR → 분류

- 스캔된 금융문서를 보험/세금/예·적금 등으로 자동 분류

- 스케줄링 → TTS

- 문서 기한 정보를 바탕으로 알림 생성

- 서류 검증

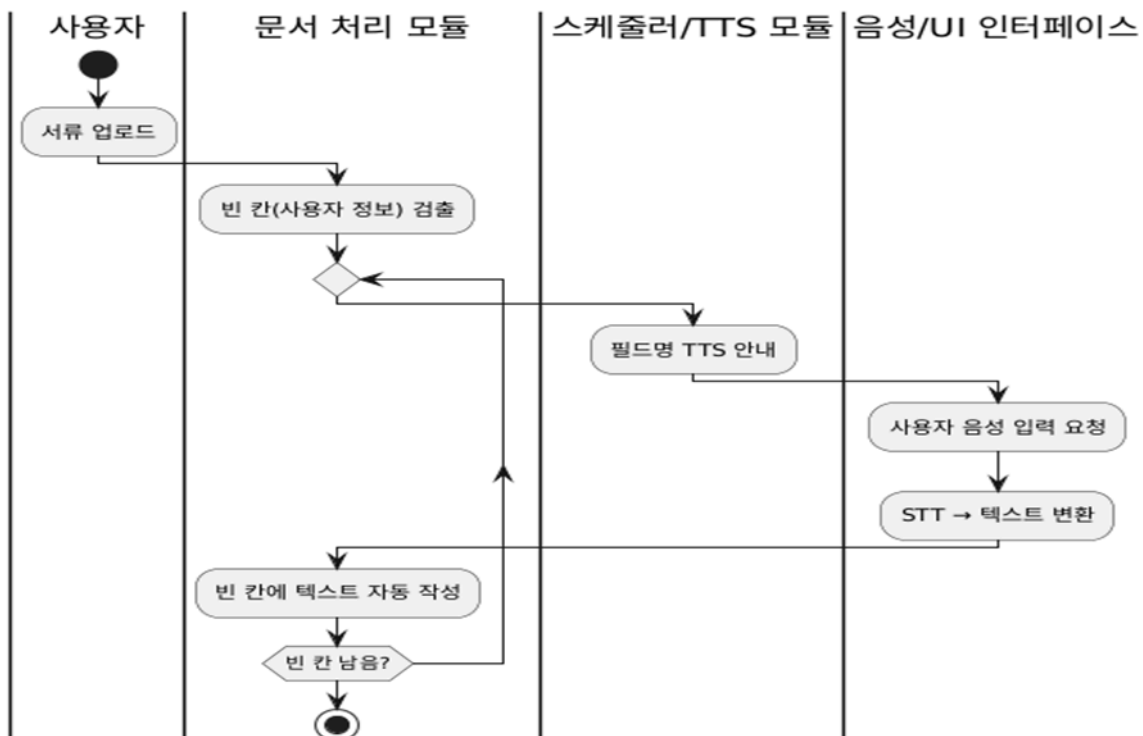
- 제출된 문서의 서류 내용 중 수치가 일치하는지 여부 검사 후 TTS 피드백

사용자는 시각 장애인을 대상으로 지정한다. 사용자는 “이체”, “문서 스캔”, “일정 알려줘” 등, 음성 명령 UI/음성 인터페이스를 통해서 음성 명령에 맞는 페이지를 띄운다. 페이지는 사용자에게 최소한의 정보를 요청한다.

이체(Transfer)

1. 사용자가 “이체” 음성 명령
2. 음성/UI 인터페이스 → 이체 페이지 오픈 → 계좌·금액 입력 요청
3. 사용자는 최소 음성 입력(예: “삼일은행 계좌 123-456, 금액 오만원”)
4. 시스템은 입력값을 화면에 표시하고, “이체를 진행하시겠습니까?” TTS 안내
5. 사용자가 “네” 음성 확인

시스템 구성도(2)



문서 스캔 및 자동 입력(Document Scan & Auto-Fill)

1. 사용자가 “문서 스캔” 음성 명령
2. 음성/UI 인터페이스 → 스캔 모드로 전환
3. 사용자가 스마트폰 카메라로 문서 촬영
4. 문서 처리 모듈
 - OCR 수행 → 텍스트 추출
 - 문서 타입(보험, 세금, 예금 등) 분류
 - 양식 내 빈 입력란(field) 자동 검출
5. 검출된 각 입력란에 대해 순차적으로 TTS 안내
 - 예: “받는 사람 이름을 말씀해주세요.”
 - 사용자는 음성으로 값을 입력(STT)
 - 시스템이 해당 필드에 자동 입력
6. 입력란 채워지면 “입력이 완료되었습니다.” TTS 안내

일정 알림(Schedule Notification)

1. 사용자가 “일정 알려줘” 음성 명령
2. 음성/UI 인터페이스 → 일정 조회 요청
3. 스케줄러/TTS 모듈
 - 사용자 계정의 캘린더 API 호출
 - 다가오는 일정(예: “만기일”) 확인
 - 필요 시, “지정일 오전 9시” 알림 스케줄 등록
 - “잔여 만기일이 3일 남았습니다.” 음성 안내
4. 사용자가 “알림 수정” 등 재요청 가능

개발 환경

개발 환경	사용 기술
딥러닝 라이브러리	Pytorch, TensorFlow
IDE	VS Code
UI(프론트 엔드)	React(TypeScript)
서버(백 엔드)	Python, fastAPI

주요 기능

입력 방식

- 시각장애인 전용 키보드 입력 지원
- 음성인식(STT) 을 통한 명령어 입력
- 복잡한 메뉴 탐색 없이 음성 명령어 한 번으로 기능 수행

문서 처리 (OCR 기반)

- 문서를 스캔하여 텍스트 데이터 추출
- 금융 분야별로 자동 분류
 - 보험, 세금, 예적금 등 카테고리별로 문서 정리
- 문서 주요 키워드 및 제목 요약 제공

일정 및 납부 알림 (TTS)

- 문서에서 기한, 만기, 납부일을 추출
- 일정 도래 시 TTS(음성 합성) 기능으로 자동 알림 제공
- 알림 주기 설정 가능

서류 검증

- 문서상의 수치(금액) 영역을 인식하여 사용자가 거래하고자 하는 수치(금액)과 일치하는지 비교
- 이상 감지 시 즉시 음성 경고 알림

구현 방식

STT (Speech-to-Text)

목적 시각장애인 사용자에게 “음성 입력” 인터페이스 제공

파이프라인

- 음성 입력 → 전처리(잡음 및 무음 제거, 로컬 모델 사용 시 MFCC 변환하여 입력)
- API 호출(Whisper / Clova / Google) → 텍스트 변환
- 필요 시 후처리 알고리즘 구현

구현 전략

- API 모델 성능 비교 (WER, 지연 시간 테스트) 후 최적의 모델을 사용.
- API 모델의 지연시간이 기준을 초과할 시 로컬 모델 튜닝하여 사용

TTS (Text-to-Speech)

목적 시각장애인 사용자에게 음성 출력 기능 제공

파이프라인

- 텍스트 입력 → 텍스트 정규화
- TTS API 호출(Naver TTS / Google TTS) → 음성 파일 생성
- 필요시 음성 후처리

구현 전략

- API 모델 성능 비교 후 최적의 모델 사용
- 지연시간 문제 발생시에 로컬 모델 사용 고려

OCR (Optical Character Recognition)

목적 금융·문서 내용의 텍스트 인식 및 접근성 지원

파이프라인

- 이미지 또는 PDF → 전처리(흑백 변환, 기울기 보정)
- API 호출(AWS Textract / Google Document AI) 또는 EasyOCR → JSON 형태 결과 획득
- 후처리(엔티티 추출, 키워드 매핑)

UI (User Interface) & Server

목적

- 시각장애인이 직관적으로 시스템을 이용할 수 있도록 접근성 보장
- 음성 기반 입력·출력, 화면 출력이 유기적으로 결합된 통합 인터페이스 제공
- 보안 및 안정성을 보장하고 API 호출 및 핵심 로직 관리

파이프라인

- React 컴포넌트 렌더링
- 사용자 음성 / 텍스트 입력 → API 호출 → API 응답 수신
- 화면 업데이트 → TTS 재생
- 피드백 루프: 재요청 / 종료

구현 전략

- React 언어를 이용해서 최소한의 입력을 STT로 요구하는 페이지 구현
- fastAPI를 이용해서 비동기 처리, 핵심 로직 관리, API 호출을 구현

4. 개발 일정 및 역할 분담

개발 일정

업무	4월	5월	6월	7월	8월	9월	10월
역할 분담 업무							
자료 수집 및 사용 환경 구성							
AI 모델 선정 및 테스트							
착수 보고서 작성 및 제출							
웹 페이지 개발							
중간 보고서 제출							
테스트 및 보완 작업							
최종 보고서 제출							

역할 분담

기획 및 데이터 수집, 전처리 : 송시우, 성민기, 성도범

음성인식 AI 모델 : 송시우

프론트엔드 개발 : 성민기

백엔드 개발 : 성도범