

상비약의 경우 약품 상자가 없다면 어떻게 카메라로 식별하나요?

= 만약 약품 상자가 없는 경우라면 개별 포장된 알약을 날개로 꺼내어 스캔할 수 있습니다.

번호별로 어떠한 정보들을 배정할 것인가요?

= 총 3개의 번호로 나누어 1번에는 약품에 관한 기본 정보(효과 포함), 2번에는 복용용량 및 방법, 3번에는 주의사항을 배정하였습니다.

크롤링

= 검색 엔진 로봇을 이용한 데이터 수집 방법. SNS, 뉴스, 웹 정보와 같이 조직 외부에서 공개된 웹문서를 수집하는 방법

대용량 연속 음성 인식

= LVCSR(Large Vocabulary Continuous Speech Recognition)

인식된 모든 음성정보 중 정의된 키워드만을 출력하게 하여, 보다 정확한 음성인식 출력 형식을 구현

SOTA (State-of-the-art)

= 현재 최고 수준의 성능을 가진 이미지 학습 모델

애플리케이션 개발

모바일 환경에서 구동하기 위해

안드로이드는 안드로이드 스튜디오 프로그램과 코틀린 언어를 사용하여 애플리케이션을 제작할 것입니다. 안드로이드 스튜디오 프로그램 안드로이드 전용 앱 제작을 위한 공식 통합개발 환경입니다. 코틀린 언어는 자바 언어를 사용할 때 사용자들이 느낀 불편한 점을 더 보완하고자 만든 프로그래밍 언어이므로 실용성이 좋아 해당 언어를 채택하게 되었습니다.

IOS는 Xcode 프로그램과 swift 언어를 사용하여 애플리케이션을 제작할 것입니다. Xcode 프로그램은 애플이 개발한 MAC OS라는 운영체제의 개발도구 모음으로 다양한 애플 제품에 대한 프로그램을 만들 수 있습니다. Swift 언어는 애플사에서 IOS와 앞서 언급했던 mac os 응용프로그램 개발을 위해 만든 프로그래밍 언어입니다. 빠른 실행 속도와 안전성 그리고 현대 프로그래밍 언어의 주요 기능들을 구현했기 때문에 선택하게 되었습니다.

TTS 시스템(타코트론2)

= 공개된 TTS 학습용 데이터인 KSS(Korean Single Speaker Speech) KSS 데이터셋을 활용하여 모델을 사전학습(pre-training)한 후 개발자의 녹음된 음성을 이용하여 fine-tuning을 진행하였습니다. 여기서 KSS 데이터셋은 전문 여성 성우 한 분이 한글 그리고 한영사전 책 4권의 예문을 읽은 약 12시간 분량의 데이터셋이며,

Fine-tuning이란 기존에 학습되어져 있는 모델을 기반으로 컴퓨터 시스템의 설계방식을 목적에 맞게 변형하고 이미 학습된 모델 가중치(Weights)로부터 학습을 업데이트하는 방법을 말합니다. KSS 모델로 사전학습을 하면 프로그램이 음성과 텍스트 사이의 정렬 규칙을 잘 학습하게 되며 이 상태에서 원하는 목소리의 녹음 파일로 학습을 하면 적은 녹음 데이터로도 양질의 TTS 파일을 만들어낼 수 있게 됩니다.

pre-training은 사전 학습이라는 의미로 정제된 데이터를 이용하여 모델을 원하는 방향으로 만들 수 있게 어느 정도 학습시켜 놓는 것.

Zeroth Project Model

= 자동음성인식 시스템에 적용되는 오픈소스 모델이며 개발자분이 만들어놓은 모델을 가져와 음성 데이터들을 추가합니다. 여기서 기본으로 100시간 정도를 훈련시키면 완성도 있는 모델이 갖춰집니다. 만약 학습이 부족하다면 데이터를 추가하여 300시간 정도를 더 훈련시키면 상용화 가능한 정도의 모델을 완성시킬 수 있습니다.