

# 음성 및 동영상 객체 인식 기반 요리 보조 시스템 개발

이종환<sup>1</sup>, 곽희웅<sup>1</sup>, 박기수<sup>1</sup>, 송미화<sup>1</sup>

<sup>1</sup>세명대학교 스마트IT학부

leejonghwan1206@naver.com, gmlndnd567@naver.com, pgs156@naver.com,  
mhsong@semyung.ac.kr

## Development of a Cooking Assistance System Based on Voice and Video Object Recognition

Jong-Hwan Lee<sup>1</sup>, Hee-Woong Kwak<sup>1</sup>, Gi-Su Park<sup>1</sup>, Mi-Hwa Song<sup>1</sup>

<sup>1</sup>School of Smart IT, Semyung University

### 요 약

모바일 서비스에서 음성인식을 활용한 애플리케이션이 가져다 주는 편리함으로 레시피 애플리케이션에 접목시켜 데이터베이스를 사용한 레시피 추천, Google Video Intelligence API를 사용하여 객체 영상분할, Google Assistant를 활용한 음성인식을 기반으로 한 레시피 애플리케이션을 제공한다.

### 1. 서론

코로나 확진자 수가 급격하게 증가하고 있는 현재, 요리는 집에서 간편하게 즐길 수 있는 취미라고 할 수 있다. 집에서 보내는 시간이 많아짐에 따라 집에서의 요리(홈 쿠킹)에 대한 관심 또한 증가하고 있다. 닐슨노르디아에서 발표한 ‘2020 대한민국 식품대전: 코로나19 이후 식품시장의 소비 트렌드’에서도 ‘외식을 줄이고 있다 (68.1%).’ 및 ‘집에서 직접 만들어서 식사한다 (58.8%).’의 응답률이 높게 나타났다 [1]. 이에 따라 홈 쿠킹을 하는 사용자들이 늘어남에 따라 레시피 제공 애플리케이션에 대한 의존도가 높아지고 있다.

기존에 제공하는 음성 요리 레시피 제공 시스템들은 대부분 사용자의 요리 숙련도에 맞추지 않고 레시피를 줄줄이 읽어주기에 급급하다. 이러한 한계가 있어, 사용자들에겐 레시피를 제어할 수 있는 요리 보조 시스템이 필요할 것으로 보인다.

이에 본 논문에서는 요리 과정에서 휴대폰을 터치할 필요 없이 레시피의 정보와 사용자의 음성으로 텍스트 및 동영상의 단계별 호출 및 제어가 가능한 시스템을 제안한다.

### 2. 관련연구

레시피를 추천하는 애플리케이션에 관한 연구는 현재 많이 개발되어 있다.

민성희 등[2]의 경우에는 레시피 데이터 기반의 식재료 궁합을 사용하여 레시피 추천 시스템을 보여준다. 본 연구에서는 구매 영수증을 영상처리 알고리즘으로 전처리 후 텍스트를 추출하여 식재료 정보를 얻어낸다. 얻어낸 텍스트 정보로 각 식재료의 궁합을 레시피 데이터로 수집하여 레시피 추천을 진행한다.

고영준[3]의 연구에서는 여러 객체 분할 방법들 중 인터랙티브 동영상 객체 분할을 중점적으로 다룬다. 사용자가 객체 분할 결과를 관찰하고, 수정하기 위한 주석을 제공한다. 여러 시나리오에 따른 분할 기술들을 제시하여 사용자가 필요에 따라 자유롭게 객체 분할을 사용할 수 있도록 제시한다.

문유경 등에서는 구글의 챗봇 플랫폼인 dialogflow를 사용하여 사용자와의 효과적인 쌍방향 소통의 시스템을 제시하였다[4]. 하지만 동영상에 대한 제어가 불가능하며, 음성인식 챗봇만을 설계해 텍스트와 동영상의 음성 제어가 동시에 이루어지지 않는다는 한계가 있다.

### 3. 구현 설계



(그림 1) 시스템 구성도

#### 3.1 음성명령

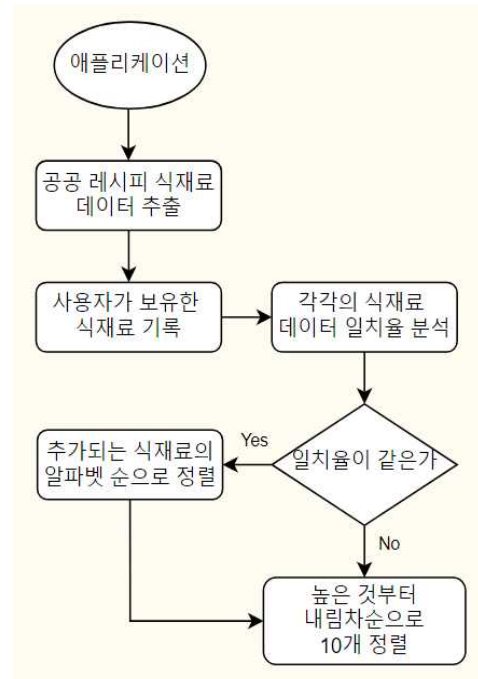
본 논문에서는 Google Assistant를 활용하여 애플리케이션에 음성명령 기능을 구현한다[5]. Google Assistant를 사용하여 애플리케이션 실행, 종료 등 사용자의 디바이스 내에서 음성명령을 통해 애플리케이션을 제어할 수 있다. 또한 애플리케이션 내에서 음성명령을 사용하여 레시피 추천, 검색, 동영상 구간 이동 등 애플리케이션의 전반적인 기능을 음성명령을 통해 실행할 수 있도록 한다.

#### 3.2 동영상 분할

동영상의 분할은 Google Video Intelligence API [6]를 이용하여 일반적으로 요리 동영상에 포함되는 특정 행위들을 인식하고 요리 동영상을 머신러닝으로 학습시켜 객체 추적 및 커스텀 항목 라벨을 생성하여 영상에서 등장하는 객체, 장소, 동작 등을 세분화하여 레시피에서 나오는 당근, 양파 등의 재료 및 썰기, 끓이기 등의 조리 동작을 객체로 인식하고, 그에 해당하는 라벨을 감지하여 요리 동영상에서 나오는 재료 또는 조리 동작으로 해당 구간을 바로 이동한다. 이러한 라벨을 바탕으로 사용자가 동영상을 시청할 때 원하는 부분을 호출하여 바로 시청할 수 있도록 한다.

#### 3.3 레시피 추천

본 논문에서 레시피 추천 기능은 사용자가 추가하여 설정하는 요리 재료에 따라 가장 많이 활용할 수 있는 요리를 추천한다.



(그림 2) 레시피 추천 알고리즘

레시피를 공공 데이터포털[7]에서 레시피 데이터를 데이터베이스에 저장하고, 레시피 재료 항목을 추출한다. 추출한 데이터는 사용자가 입력한 데이터와 대조하여 일치율을 검사한다.

(a)	(b)	(c)
소세지 1개	소세지 1개	소세지 1개
고추장 2T	고추장 2T	고추장 2T
설탕 1T	설탕 1T	설탕 1T
고춧가루 2T	고춧가루 2T	고춧가루 2T
간장 3T	간장 3T	간장 3T
양파 1개	양파 1개	양파 1개
파 1/2개	파 1/2개	파 1/2개
	돼지고기 300g	돼지고기 300g
		라면사리 1개

(그림 3) 레시피 데이터 예시

본 논문에서는 일치율을 계산하고 비교하기 위해 자카드 유사도[8]를 이용한다. 자카드 유사도란 두 집합이 동일하면 1의 값을 가지고, 공통된 원소가 없다면 0의 값을 가진다. 다음 수식 (1)의 A는 그림 3의 (a), B는 (b)라고 가정한다. 교집합으로 표현된 겹치는 식재료 수는 7개이고 합집합으로 표현된 총 식재료 수에서 겹치는 식재료 수를 뺀 값은 8이다. 결과적으로 (a)와 (b)의 유사도는 0.875이고, 똑같은 방법으로 계산한 (a)와 (c)의 유사도는 약 0.78이 나온다. 즉, 사용자가 입력한 식재료가 완전히 들어있는 레시피 일치라도 유사도가 다르다. 하지만 추가되는 식재료가 1개이지만 서로 다른 식재료라면 일

치율이 같아도 다른 레시피가 되는데, 이것을 방지하기 위해 식재료 데이터 일치율을 분석한 후에 곧바로 일치율이 같은 레시피가 있다면 추가되는 식재료의 알파벳 순으로 정렬한다. 즉, 사용자는 레시피 추천 알고리즘이 일치율을 검사하고 나열한 정보를 최종적으로 보게 된다.

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$

(1)

#### 4. 결론

본 논문에서는 음성 제어와 동영상 객체 인식 및 라벨 감지를 이용하여 화면에서 추가로 사용자의 직접적인 조작 없이 사용자가 원하는 구간으로 이동할 수 있도록 하였고, 보유 재료에 가장 일치율이 높은 레시피를 추천하는 알고리즘을 더해 사용자의 전반적인 요리 과정을 보조하는 시스템을 제안하였다.

기존의 요리 보조 시스템에서 제공하는 동영상 객체를 인식하고 라벨을 감지하여 동영상의 해당 부분을 찾아가는 기능에 대한 구현이 되어 있지 않아 본 논문에서는 시스템을 개선 및 구현하기 위해 Google Video Intelligence의 객체 추적 및 라벨 감지 기능을 이용해 동영상에서 원하는 구간을 지연 시간 없이 찾을 수 있도록 설계하였다. 수집한 레시피 데이터에 포함되어 있는 재료 데이터를 사용자의 입력 데이터와 비교해 일치율을 검사하고, 일치율이 같을 때에는 세부 알고리즘을 통해 사용자에게 노출되는 추천 레시피의 순서를 정하였다.

향후 연구로는 동영상의 객체 및 명사를 찾아가는 것에서 끝나는 것이 아니라, 음성 인식으로 사용자의 편의를 더해줄 수 있는 기능을 개발하고, 레시피의 세부 내용들에 대해 종합적인 음성 제어를 위한 기능이 필요하다. 추천 시스템의 경우 사용자가 직접 데이터를 입력하는 것이 아니라 자동화 할 수 있는 방법과 보유 재료의 세부 현황을 이용해 더욱 정확한 일치율을 구체화하는 알고리즘을 개발할 것이다.

#### 참고문헌

- [1] Ji Young Kwon, Sang-Wook Song, Ha-Na Kim, Sung Gu Kang, "Changes in Body Mass Index and Prevalence of Metabolic Syndrome during COVID-19 Lockdown Period," pp.208, 2021
- [2] Y. S. Oh and S. H. Min, "Implementation of Recipe Recommendation System Using Ingredients Combination Analysis based on Recipe Data", *Journal of Korea Multimedia Society*, Vol. 24, No. 8, (pp. 1114-1121) August 2021
- [3] 고영준, "딥러닝 기반 동영상 객체 분할 기술", *전자공학회지*, Vol. 48, No. 8, (pp. 50-55) 2021
- [4] Moon, Y.-G., Kim, G.-Y., Kim, Y.-H., Park, M.-J., Seo, M.-H., & Nah, J.-E.. "Development of an interactive smart cooking service system using behavior and voice recognition." *Proceedings of the Korea Information Processing Society Conference*, pp. 1128 - 1131 2021
- [5] Google Assistant(2022), "https://developers.google.com/assistant/app/get-started" (accessed April 09, 2022)
- [6] Google Video Intelligence(2022), "https://cloud.google.com/video-intelligence?hl=ko" (accessed April 13, 2022)
- [7] Recipe Standard Information(2022), <https://www.data.go.kr/data/15057205/openapi.do> (accessed April 13, 2022).
- [8] Jaccard similarity(2022), [https://ko.wikipedia.org/wiki/%EC%9E%90%EC%B9%B4%EB%93%9C\\_%EC%A7%80%EC%88%98](https://ko.wikipedia.org/wiki/%EC%9E%90%EC%B9%B4%EB%93%9C_%EC%A7%80%EC%88%98) (accessed April 13, 2022).