

이 게시물의 토론, 통계 및 저자 프로필은 <https://www.researchgate.net/publication/356814950> 에서 확인하세요.

## 딥러닝을 활용한 재료 감지 기반 음식 레시피 추천

컨퍼런스 논문 - 2021년 12월

인용

1

읽기

2,683

2명의 저자를 포함합니다:



Md. 키쇼르 모룰 코넬

대학교

38개 출판물 55회 인용

[프로필 보기](#)

이 간행물의 저자 중 일부는 이러한 관련 프로젝트에도 참여하고 있습니다:



스피치 [뷰 프로젝트](#)를 통한 화자의 성별 인식 및 나이 추정 [프로젝트](#)



혈청 비타민 D 수치와 파킨슨병의 연관성 [프로젝트 보기](#)

이 페이지의 모든 콘텐츠는 [Md. Kishor Morol](#)에 의해 2022년 3월 17일에 업로드되었습니다.

사용자가 다운로드한 파일에 대한 개선을 요청했습니다.

# 딥러닝을 활용한 재료 감지 기반 음식 레시피 추천

Md. 샤파트 자밀 로콘, 키쇼르 모롤, 이스라 빈테 하산, A. M. 사이프, 라피드 후세인 칸  
컴퓨터 과학과

아메리칸 국제 대학교-방글라데시(AIUB)

{shafaatjamilrokon, kishoremorol, ishrahaman12, alimdsaif1996, rafidkhan321.rk}@gmail.com

## 초록

음식은 인간의 생존에 필수적인 요소이며, 사람들은 항상 다양한 종류의 맛있는 레시피를 맛보려고 노력합니다. 사람들은 종종 식재료의 이름도 모른 채 식재료를 선택하거나 식료품점에서 잘 모르는 식재료를 구입하는 경우가 많습니다. 맛있는 음식 레시피를 만들기 위해 어떤 재료를 섞을 수 있는지 아는 것은 필수입니다. 재료 목록을 보고 적합한 레시피를 선택하는 것은 초보 요리사에게는 매우 어려운 일입니다. 하지만 전문가에게도 문제가 될 수 있습니다. 일상 생활에서 머신러닝은 끊임없이 사용되고 있습니다. 이미지 처리를 통해 사물을 인식하는 것이 그러한 예 중 하나입니다. 이 과정은 음식 재료가 다양하기 때문에 복잡하지만, 기존의 접근 방식은 부정확한 결과를 초래합니다. 이러한 문제는 머신러닝과 딥러닝 접근 방식을 통해 해결할 수 있습니다. 이 논문에서는 식재료 인식 모델을 구현하고 인식된 식재료를 기반으로 레시피를 추천하는 알고리즘을 설계했습니다. 이를 위해 32가지 식재료 클래스에 속하는 9856개의 이미지로 구성된 맞춤형 데이터셋을 만들었습니다. 식재료 인식에는 컨볼루션 신경망(CNN) 모델을 사용했고, 레시피 추천에는 머신러닝을 활용했습니다. 94%의 정확도를 달성했는데, 이는 매우 인상적인 결과입니다.

드 후세인 칸. 2022. 딥러닝을 이용한 재료 감지에 기반한 음식 레시피 추천. 2022년 3월 10-12일, 방글라데시 다카, 국제 컴퓨팅 발전 컨퍼런스(ICCA 2022)에서.

## 키워드

딥러닝, CNN, 재료 감지, 레시피 추천, 리셋50

## ACM 참조 형식:

Md. 샤파트 자밀 로콘, 키쇼르 모롤, 이스라 빈테 하산, A. M. 사이프, 라피

## 1 소개

요즘 사람들은 건강에 대한 관심이 매우 높아져 식사에 최대한 건강한 음식을 섭취하려고 노력합니다. 적절한 음식 레시피는 음식의 진정한 맛과 균형 잡힌 식단, 그리고 건강한 삶을 제공해야 합니다. 우리 모두 알다시피 건강은 재산이므로 더 나은 건강을 위해 음식 선택은 필수적입니다. 때때로 사람들은 이름도 모르는 재료를 선택하기도 합니다. 또는 식료품점에서 식재료를 구입했지만 해당 식재료를 사용하여 레시피를 만드는 방법을 모릅니다. 사람들은 어떤 재료를 섞어서 맛있는 레시피를 만들 수 있는지 알아야 합니다. 초보자의 경우 재료만 보고 레시피를 선택하는 것은 매우 어렵습니다. 전문 요리사에게도 어려운 일입니다. 이 논문에서는 이 문제를 해결하고자 합니다. 식재료를 인식하고, 이렇게 인식된 식재료를 분석 및 분류한 후 그에 맞는 레시피로 요리할 수 있는 음식을 추천하는 방식을 구현했습니다. 이 추천은 사람들이 적합한 음식 레시피를 얻는 데 도움이 될 것입니다. 온라인 요리 레시피 사이트는 매우 인기가 있지만, 사람들이 이러한 사이트를 따라 요리하려고 할 때 때때로 혼란을 겪습니다. 이로 인해 사용자가 식료품점에서 검색하는 동안 이러한 유형의 요리 사이트를 참조하지 않을 수 있습니다. 이러한 문제를 해결하기 위해 기술을 개선해야 했고, 그 결과 사물 인식 기술이 세상에 소개되었습니다. 사물 인식 기술에는 상당한 발전이 있었습니다. 객체 인식은 이미지에 표시된 객체의 특정 속성을 기반으로 객체를 식별하는 기술입니다[1]. 즉, 객체 인식은 이미지를 분석하여 현실 세계의 사물을 찾아내는 과정입니다. 객체 인식, 컴퓨터 비전, 딥러닝 모델을 사용하기 위해 오픈 컴퓨터 비전 라이브러리(OpenCV)[2], 텐서플로우[3], 넘파이[4], 케라스[5]와 같은 오픈소스 소프트웨어 라이브러리가 널리 사용되고 있습니다. 이러한 라이브러리를 사용하면 PC와 아이폰, 안드로이드 스마트폰과 같은 모바일 기기에서 객체 인식 체계를 구현할 수 있습니다. 최근 모바일 기기의 발전과 폭발적인 보급으로 모바일 기기에서 언제든지 사물을 인식할 수 있게 되었습니다. 이미지 처리 및 표현에는 OpenCV[2]를 사용할 수 있습니다. NumPy와 같은 딥러닝 API는 수치 연산을 수행할 수 있어 머신러닝을 더 빠르고 쉽게 수행할 수 있습니다. TensorFlow[3] API는 이미지를 인식하고 이미지 분류를 위한 딥 러닝 아키텍처를 훈련하는 데 사용할 수 있습니다.

MobileNet [6], ResNet [7], Inception [8]과 같은 VGG16 [9]도 사용할 수 있습니다. 사전 학습된 아키텍처를 사용하면 전이 학습 [10]을 통해 아키텍처를 수정할 수 있다는 장점이 있습니다. 이러한 라이브러리, API 및 아키텍처의 도움으로 이미지에서 객체를 쉽게 인식할 수 있습니다.

레시피 식별의 경우 제대로 된 데이터 세트가 없어 연구가 거의 이루어지지 않았고, 여러 객체를 인식하는 작업도 여전히 어렵습니다. 저희가 찾은 대부분의 논문은 레시피 식별이나 음식 추천에 관한 것이었습니다. 하지만 저희는 분류율을 최대로 끌어올려 식 재료를 검출하고 가치 있는 음식 레시피를 추천하는 것을 목표로 합니다. 다음 섹션에서는 관련 선행 연구에 대해 설명합니다.

## 2 관련 작품

음식 감지, 재료 감지 또는 레시피 추천은 최근 몇 년 동안 점점 더 많은 관심을 받고 있습니다. 이러한 모든 딥러닝 기반 작업은 객체 감지 및 분류와 관련이 있습니다. 객체 감지 및 분류를 위한 적절한 딥러닝 모델과 방법이 존재합니다. 컬러 히스토그램, BoF, 선형 커널 SVM 분류기, K-최근이웃 등을 사용하는 것이 잘 알려져 있습니다.

Kawano 등[12]은 스마트폰용 실시간 음식 인식 시스템에 대한 접근 방식을 제시했습니다. 두 가지 유형의 실시간 이미지 인식 방법이 사용되었습니다. 하나는 특징 가방(BoF)과 컬러 히스토그램을 X2 커널 특징 맵과 결합한 것입니다. 두 번째는 최첨단 피쳐 벡터 표현을 사용하는 HOG 패치 디스크립터와 컬러 패치 디스크립터입니다. 분류자로는 선형 SVM이 사용되었습니다. 79.2%의 분류 정확도를 달성했습니다.

볼라노스 등[13]은 음식 이미지 인식에 CNN을 사용했습니다. 이 방법에는 두 가지 입력을 사용합니다. 첫 번째는 음식 이미지에 대한 낮은 수준의 설명을 제공합니다. 첫 번째 방법에서는 InceptionResNetV2 CNN의 두 번째 계층이 사용되며, 두 번째 계층은 LogMeals API를 사용하여 음식 이미지에 대한 높은 수준의 설명을 제공합니다. 이 LogMeals API는 음식 그룹, 재료, 요리를 예측하는 세 가지 CNN을 제공했습니다.

Chang Liu 등[14]은 시각 기반 음식 이미지 인식을 위한 CNN 기

반의 새로운 접근법을 7 계층 아키텍처로 제안하고 UEC-100 데이터 세트를 사용하여 93.7%의 상위 5 개 정확도를 달성했으며, 기존 접근 방식인 SURF- BoF + 컬러 히스토그램 및 MKL은 동일한 데이터 세트로 68.3 % ~ 76.8 %의 상위 5 개 정확도를 달성할 수 있었습니다.

Raboy-McGowan과 L. 곤잘레스 등[15][16]은 레시피 1M 데이터셋을 사용했으며, 레시피 1M 데이터셋에서 학습된 음식 대 레시피 생성기인 레시피 넷(Recipe Net)을 도입했습니다. 이들은 음식 이미지를 분류하고 특징을 인코딩하기 위해 ResNet-50, DenseNet-121 컨볼루션 신경망을 사용했습니다. 그 후, 그들은 K-

를 사용하여 가장 가까운 이웃에게 레시피 1백만 개 데이터 세트의 레시피를 추천할 수 있습니다.

Chen, J. 외[17]는 Go Cooking 웹사이트에서 수집한 61,139개의 이미지-레시피 쌍으로 구성된 데이터 세트를 사용했습니다. 이들은 앞서 언급한 문제를 해결하기 위해 크로스 모달리티 학습의 최근 발전을 탐구했습니다. 딥 모델 스택 주의 네트워크(SAN)를 시스템에 흡수하고 수정했습니다.

KeijiYanai 등[18]은 안드로이드 OS용 식품 성분 인식 시스템을 만들었습니다. 이 시스템은 이미지 파일로 여러 프레임에서 추출한 컬러 히스토그램 기반 특징 벡을 사용하고 분류기로 선형 커널 SVM을 사용했습니다. 30개의 동영상 데이터를 데이터베이스에 저장했고 정확도는 83.93%였습니다.

수야시 마헤슈와리(Suyash Maheshwari) 등[19]은 식재료를 추천하기 위해 두 가지 알고리즘을 제안했습니다. 요리 및 시식 실험을 통해 제안된 방법들이 각각의 목적에 효과적임을 확인했습니다. 추천된 대체 식재료를 사용한 요리 시연과 주관적 평가 실험을 통해 두 알고리즘 모두 90% 이상 허용 가능한 식재료를 추천하는 것으로 확인되었습니다. 다만, 알고리즘 1은 대체 식재료와 유사한 대체 식재료를 추천하지 않는 경우가 있었습니다.

모나 미슈라, 이판 공 등[20]의 연구에서는 이미지 인식을 위한 CNN을 구축하기 위해 다양한 카테고리에 대해 수천 개의 이미지를 학습시킵니다. 그리고 사용자가 업로드한 이미지 속 사물이 무엇인지 예측합니다.

따라서 우리는 식재료를 식별하고 이를 기반으로 레시피를 추천할 수 있는 모델을 만들 수 있는 충분한 범위가 있다는 것을 이해했습니다. 다음 섹션에서는 우리가 제안한 모델에 대해 설명하겠습니다.

### 3 방법론

#### 3.1 제안 방법

먼저 작업을 두 부분으로 나누었습니다. 하나는 컨볼루션 신경망(CNN)을 사용하여 32가지 재료를 식별하고 결과를 얻은 후 두 번째 부분으로 진행합니다. 두 번째 파트는 19가지 클래스가 포함된 레시피 추천입니다.

첫 번째 부분에서는 CNN 모델 훈련에 전이 학습을 사용했습니다. 사전 학습된 모델을 재사용하는 것을 전이 학습이라고 합니다. 사전 훈련된 모델로는 ResNet50을 사용했습니다. 전이 학습은 비교적 적은 데이터로 심층 신경망을 훈련할 수 있기 때문에 딥 러닝에서 매우 인기가 있습니다[11]. 대부분의 실제 문제에는 복잡한 심층 신경망을 훈련하기 위한 수백만 개의 레이블이 지정된 데이터 포인트가 없기 때문에 전이 학습이 효율적입니다.

기본 모델로는 50개의 레이어로 구성된 심층 컨볼루션 신경망인 ResNet50을 사용했습니다. ResNet50은 다음과 같은 학습을 거쳤습니다.

이미지넷 데이터 세트에 백만 개의 이미지가 있습니다. 사전 학습된 네트워크는 이미지를 1000개의 카테고리로 분류할 수 있지만, 볶음밥, 초트뽕티, 지하철 등과 같은 32개의 객체 카테고리가 필요합니다. 사전 학습된 ResNet50은 이러한 32개의 특정 카테고리로 이미지를 분류하지 않습니다. 이 특정 목적을 위해 처음부터 새로운 모델을 만들었지만 좋은 결과를 얻으려면 볶음밥과 초트뽕티 등을 판별할 수 있는 레이블이 있는 이미지가 많이 필요합니다.

사용할 수 있으며 예측을 생성하는 데 사용되는 최종 레이어만 교체하면 됩니다. 그렇기 때문에 1000개 클래스에 대해 학습된 ResNet50 모델의 마지막 예측 레이어를 줄이고

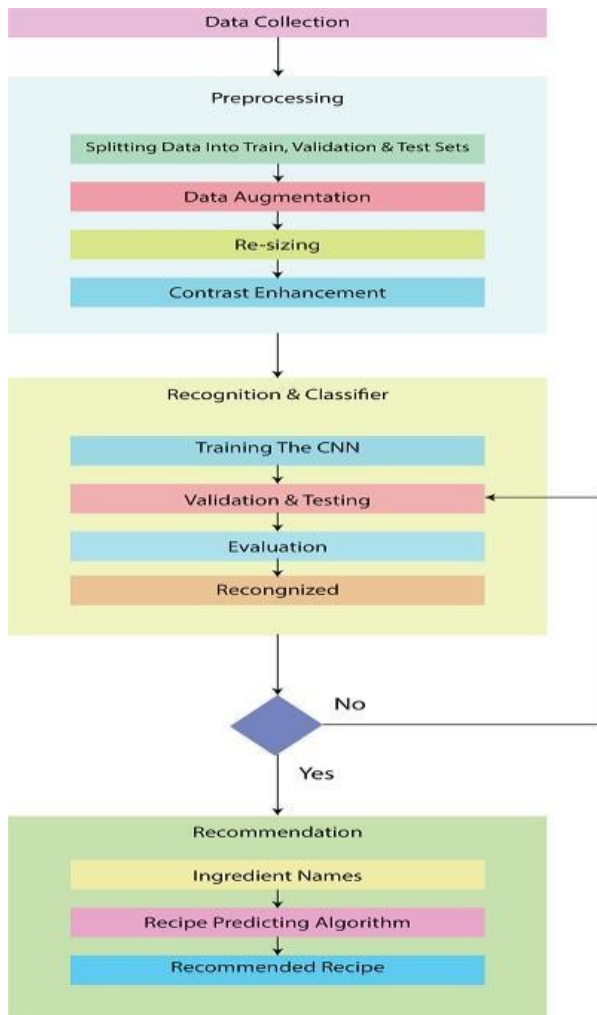


그림 1. 제안된 방법의 블록 다이어그램

전이 학습을 사용하면 더 적은 데이터로도 좋은 결과를 얻을 수 있었습니다. 이것이 바로 전이 학습을 선택한 이유입니다. 딥러닝 모델의 초기 레이어는 모양을 식별하고, 마지막 레이어는 더 복잡한 시각적 패턴을 식별하며, 마지막 레이어는 예측을 수행합니다. 대부분의 컴퓨터 비전 문제와 관련된 저수준 패턴이 유사하기 때문에 사전 학습된 모델의 대부분의 레이어는 새로운 애플리케이션에 유용합니다. 즉, 사전 학습된 ResNet50 모델의 레이어 대부분을 재

플랫화 레이어와 필요한 32개 클래스에 대한 새로운 예측 레이어를 추가했습니다.

그림 2. 제안된 전이 학습 모델의 블록 다이어그램

사전 학습된 모델에서 그 이전의 일부 레이어는 도로, 건물, 창문 등과 같은 특징을 식별할 수 있으며, 이 새로운 마지막 레이어는 ResNet50 모델의 마지막 레이어를 대체하여 이미지가 쇼트포티인지 지하철인지 또는 30개의 클래스 중 다른 것인지를 예측합니다.

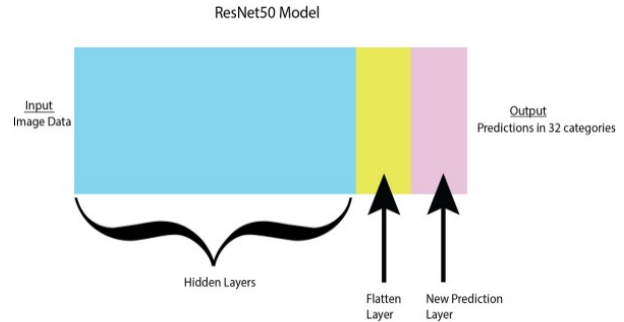
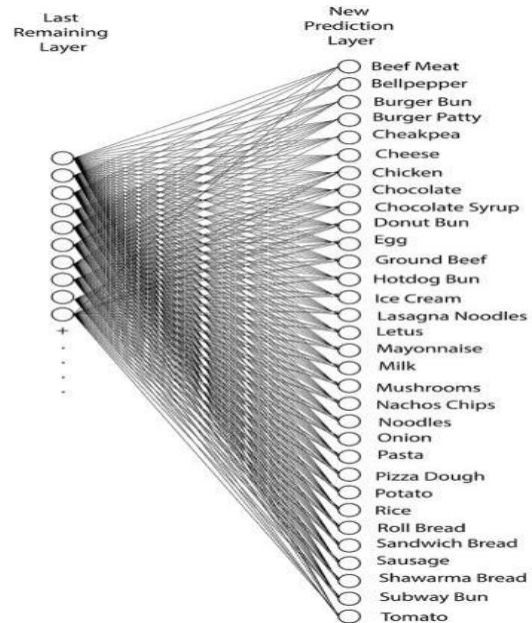


그림 3. 마지막 남은 레이어와 새로운 예측 레이어 간의 연결

여기에는 많은 연결이 있습니다. 그림 3에서 ResNet50 모델에는 많은 레이어가 있으며, 마지막 레이어를 잘라낸 것을 볼 수 있습니다. 마지막 레이어에 남은 것은 텐서에서 일련의 숫자로 저장된 이미지 콘텐츠에 대한 정보입니다. 이 텐서는 벡터라고도 하는 1차원 텐서여야 합니다. 일련의 점으로 표시할 수 있습니다. 점을 노드라고 합니다. 첫 번째 노드는 벡터에서 첫 번째 숫자를 나타내고 두 번째 노드는





두 번째 숫자 등입니다. 이미지를 32개의 카테고리로 분류하려고 합니다: 소고기, 피망, 버거 번, 버거 패티, 병아리콩, 치즈, 닭고기, 초콜릿, 초콜릿 시럽, 도넛 번, 계란, 같은 소고기, 핫도그 번, 아이스크림, 라자냐면, 레투스, 마요네즈, 우유, 버섯, 나초칩, 국수, 양파, 파스타, 피자 반죽, 감자, 쌀, 롤빵, 샌드위치 빵, 소시지, 사와르마 빵, 서브웨이 번 및 토마토입니다. 마지막 레이어에서는 사전 학습된 모델을 유지하고 32개의 노드가 있는 다른 레이어를 추가합니다. 첫 번째 노드는 소고기 고기 이미지를, 두 번째 노드는 피망 이미지를, 세 번째 노드는 햄버거 빵 이미지를 캡처하는 방식으로 학습합니다. 학습 데이터를 사용하여 어떤 노드가 어떤 이미지가 버거 번인지, 어떤 노드가 파스타인지 등을 결정합니다.

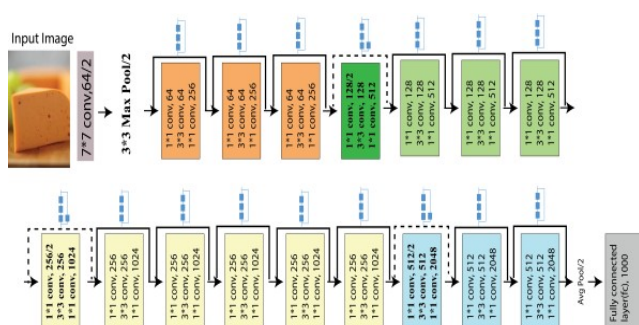
### 3.2 아키텍처

ResNet [7]의 변형입니다. ResNet 50에는 48개의 컨볼루션 레이어와 1개의 MaxPool 및 1개의 평균 풀 레이어가 있습니다. 이 네트워크에는 총 50개의 레이어가 있습니다.

- 첫 번째 레이어의 피쳐 크기는  $7 \times 7$ 이고 이러한 필터 64개는 모두 크기 2의 보폭을 갖습니다. 그리고 하나의 레이어를 제공합니다.
- 다음으로 창 크기가  $3 \times 3$ 이고 보폭 크기가 2인 최대 풀링을 볼 수 있습니다.
- $1 \times 1$ , 64는 첫 번째 컨볼루션 레이어의 필터 크기가  $1 \times 1$ 이고 이러한 필터가 64개라는 의미입니다. 다음  $3 \times 3$ , 64는 다음을 의미합니다.  
필터 크기는  $3 \times 3$ , 이러한 필터는 64개, 마지막으로  $1 \times 1$ 입니다,  
256은  $1 \times 1$ 의 필터 크기와 256개의 필터를 의미합니다.  
이 세 가지 레이어가 3번 반복되므로 이 단계에서는 9개의 레이어가 사용됩니다.

그림 4. ResNet50의 아키텍처

- 그 후  $1 \times 1$ , 128 레이어,  $3 \times 3$ , 128, 마지막  $1 \times 1$ , 512에서 이 단계를 4번 반복하므로 이 단계에서 12개의 레이어가 생깁니다.
- 다음으로  $1 \times 1$ , 256 레이어,  $3 \times 3$ , 256,  $1 \times 1$ , 1024를 6번 반복하여 18개의 레이어를 만듭니다.
- 그리고 다시  $1 \times 1$ , 512 레이어,  $3 \times 3$ , 512 및  $1 \times 1$ , 2048을 3번 반복하여 9개의 레이어를 만들었습니다.



[illegible]

개발한 알고리즘은 이 데이터베이스에서 선형 검색을 수행하여 닭고기, 달걀, 쌀 식재료 옆에 '1'이 할당된 레시피와 다른 식재료 옆에 '0'이 할당된 레시피를 교차 확인합니다. 그러면 볶음밥을 찾을 수 있습니다. 자체 개발한 알고리즘을 통해 여러 레시피를 추천할 수 있습니다. 예를 들어 인식된 식재료는 병아리콩, 달걀, 감자입니다. 먼저 레시피 추천 알고리즘은 데이터베이스에서 병아리콩, 달걀, 감자에 대해 개별적으로 선형 검색을 수행하고 병아리콩에만 "1"이 할당된 레시피와 달걀 또는 감자에만 "1"이 할당된 레시피를 교차 확인한 다음 병아리콩, 달걀, 감자에 대해 모두 "1"이 할당된 레시피를 교차 확인합니다. 그러면 휘귀, 감자튀김, 포테이토 웨지가 표시됩니다. 이것이 저희가 개발한 레시피 추천 알고리즘이 작동하는 방식입니다.

### 3.3 데이터 세트

피자 도우, 양상추, 마요네즈 등 이 주제에 대한 실제 데이터셋이 온라인에 존재하지 않았기 때문에 직접 데이터셋을 만들어야 했습니다. 이 연구에서는 Kaggle의 세 가지 데이터 세트인 food101, fruit 360, UECFOOD256과 우리가 직접 만든 데이터를 사용했습니다. 과일360에서는 감자와 토마토 데이터셋을, food101에서는 아이스크림과 양파 데이터셋을, 마지막으로 UECFOOD256에서는 쌀, 치킨, 롤빵, 소시지 데이터셋을 가져왔습니다. 총 32가지 재료에 대한 9856개의 이미지가 포함된 데이터셋을 그림 6과 같이 만들었습니다.



그림 6. 데이터 세트에 포함된 32종의 식품 성분

### 3.4 전처리

컨볼루션 신경망은 동일한 크기의 입력을 받아들이기 때문에 모든 식재료 이미지는 CNN에 입력하기 전에 식별 가능한 크기로 조정

해야 합니다. ResNet50의 목표 입력 크기는  $224 \times 224$ 였습니다. 하지만 데이터 세트 이미지는 다양한 출처에서 수집했기 때문에 모양과 크기가 모두 달랐습니다. 일부 식재료 데이터 세트는 Kaggle에서 무료로 사용할 수 있었고, 다른 데이터 세트는 카메라로 직접 만든 것이었습니다. 변형되지 않은 이미지로 CNN을 훈련하면 훌륭한 분류 전시회가 열릴 것입니다. 그래서 이미지의 크기를  $224 \times 224$ 픽셀로 조정하고 이미지에서 불필요한 물체를 제외했습니다.

### 3.5 데이터 증강

머신러닝 모델은 데이터 세트가 풍부하고 적절할 때 더 나은 성능을 발휘하고 더 정확합니다. 모든 재료 중 양파, 토마토, 감자, 쌀에 대한 데이터 세트가 미리 준비되어 있습니다. 소고기, 피망, 버거 번, 버거 패티, 병아리콩, 치즈, 닭고기, 초콜릿 시럽, 도넛 번, 같은 소고기의 사진 수를 늘리기 위해 데이터 증강을 사용했습니다, 핫도그 번, 라자냐면, 양상추, 마요네즈, 우유, 버섯, 나초칩, 국수, 파스타, 계란, 피자 도우, 초콜릿, 롤빵, 샌드위치 빵, 소시지, 사과르마 빵, 서브웨이 번. 데이터 증강을 수행하기 위해 텐서플로우 라이브러리를 사용했습니다. 이미지 데이터 생성기 함수는 Keras 함수인 데이터를 보완하는 데 사용되었습니다. 데이터 증강 과정에서 이미지를 45도 회전하고, 폭과 높이를 20% 이동하고, 20% 확대하고, 가로로 뒤집고, 20% 범위로 잘라내는 작업을 수행했습니다. 이미지 데이터 증강은 일반적으로 검증 또는 평가 데이터 세트가 아닌 훈련 데이터 세트에만 적용됩니다.

### 3.6 옵티마이저 및 학습 속도

최적화에서 옵티마이저는 손실을 줄이기 위해 가중치 및 학습 속도와 같은 심층 신경망 속성을 수정하는 알고리즘 또는 방법입니다. 옵티마이저는 최적화 문제 해결의 일부로 함수를 최소화하는 데 사용됩니다. 최적화 기법에는 RMSprop, Adam, Nadam, SGD, Ftrl 등 다양한 종류가 있습니다. 본 연구에서는 아담 최적화 기법[21]을 사용했습니다. 아담 최적화는 컴퓨터 비전과 자연어 처리에 사용하기 위해 확률적 경사 하강을 확장하는 기법입니다. 더 나은 성능으로 인해 PC 비전 연구자들이 일반적으로 사용합니다. 저희 모델에서는 아담 최적화 기법을 0.001의 학습률로 사용했습니다. 이것이 아담 옵티마이저의 기본 학습 속도입니다. 컨볼루션 신경망을 훈련할 때 학습률은 중요한 역할을 합니다. 학습률이 낮을수록 더 나은 분류 결과를 얻을 수 있지만, 최적화기는 글로벌 최적값을 얻기 위해 더 많은 노력을 기울여 손실을 줄입니다. 학습률이 높을수록

정확도가 떨어집니다. 학습률은 유효성 검사 정확도를 확인하면 자연스럽게 떨어지기 때문에 0.001을 학습률로 설정했습니다. 손실을 계산하기 위해 범주형 교차 엔트로피 함수를 사용했습니다. 범주형 교차 엔트로피는 소프트맥스 손실이라고도 합니다. 이것은 교차 엔트로피 손실이 있는 소프트맥스 활성화입니다. 분류 오차, 평균 제곱 오차 등과 같은 다른 함수와 비교했을 때 교차 엔트로피가 좋은 품질을 보인다는 것을 발견했습니다[22] 우리 모델은 다중 클래스 분류 문제이므로 범주형 교차 엔트로피가 가장 적합한 선택입니다.

## 4 결과 및 토론

### 4.1 모델 성능

앞서 언급했듯이 총 9856개의 이미지로 구성된 데이터 세트가 있습니다. 따라서 이미지의 70%는 학습용으로, 20%는 테스트용으로, 나머지 10%는 검증용으로 사용했습니다. 전이 학습을 사용하여 20개의 에포크를 실행한 결과, 제안한 CNN 기반 ResNet50 모델은 훈련 데이터 세트에 대해 99.71%의 정확도를 달성했습니다. 그리고 검증 데이터 세트에서는 92.6%의 정확도를 기록했습니다. 결과와 혼동 행렬을 분석한 결과, CNN 모델은 식품 성분 분류에 적합한 성능을 가지고 있다고 말할 수 있습니다. CNN 모델의 성능은 아래 그림 7과 8에 나와 있습니다.

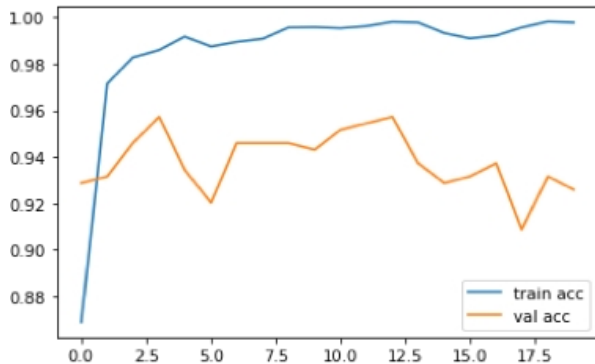


그림 7. 훈련 및 검증 정확도.



그림 8. 교육 및 검증 손실

### 4.2 결과 비교

레시피 추천에 대한 연구는 거의 찾아볼 수 없습니다. 저희 연구와 비슷한 음식 인식에 관한 논문이 몇 편 있습니다.

표 1. 이전 작품들 간의 비교

업무	기술 사용됨	정확도
Khan 외, "정크푸드를 인식하기 위한 머신러닝 접근 방식 (제안된 모델)" [23]	CNN	90.47%
Sun, 외. "이미지 특징 추출을 통한 Yelp 음식 식별 및 분류" [24]	CNN & SVM	68.49%
"사전 학습 및 미세 조정을 통해 딥 컨 볼루션 네트워크를 사용한 음식 이미 지 인식" [25]	DCNN	70.4%
박스터 외. "식품 인식을 이용한 성분 수준 기능" [26]	CNN	81.2%
R. Xu 등 "모델링 레스토랑 음식 인식을 위한 컨텍스트" [27]	CNN	84.9%
Ciocca 외. "음식 인식 : 새로운 데 이터 세트 실험, 그 리고 결과" [28]	CNN	79%
메이어스 외. "Im2칼로리: towards 자동화된 모바일 비전 푸드 다이어리" [29]	구글러 넷	79%
<b>딤러닝을 사용한 재료 감지를 기반으로 한 레시피 추천 학습</b>	<b>CNN 기반 Resnet50</b>	<b>94%</b>

## 4 결론

본 논문에서는 식재료를 인식하는 CNN 모델과 감지된 식재료를 기반으로 요리 레시피를 제안하는 레시피 추천 알고리즘을 제시했습니다. 또한 32개 카테고리의 식재료로 구성된 맞춤형 데이터셋을 도입했습니다. 테스트 결과 94%라는 높은 정확도를 달성

하며 이미지에서 식재료를 인식하는 모델의 성능이 한층 고도화되었음을 입증했습니다. 이러한 높은 정확도 덕분에 CNN이 식재료 인식에 특히 적합하다는 것을 알 수 있었습니다. 또한, 앞서 언급했듯이 컴퓨터 GPU의 성능이 높아진다면 정확도는 현재의 정확도 수준보다 더 높아질 수 있을 것으로 판단하여 CNN 모델 학습에 전이 학습을 사용했습니다. ResNet50을 기본 모델로 사용했습니다. 실험의 일환으로 VGG16과 MobileNetV2도 기본 모델로 사용했지만, ResNet50이 더 나은 결과를 생성했습니다.

사전 학습된 다른 두 모델보다 성능이 뛰어납니다. 32가지 식재료에 대해 19가지 요리 레시피를 선정하고 식재료 감지를 기반으로 레시피 추천을 위한 고유한 알고리즘을 개발했습니다.

앞으로 레시피와 식재료의 수를 늘리고 데이터셋을 더욱 풍부하게 만드는 것을 목표로 하고 있습니다. 데이터 세트에서 두 가지 문제를 발견했는데, 하나는 이미지의 대비 변화이고 다른 하나는 클래스의 불균형 문제입니다. 불균형 문제를 줄이기 위해 데이터 증강을 사용했습니다. 앞으로는 이 두 가지 문제를 완전히 극복하기 위해 노력할 것입니다. 우리 연구에는 한 가지 한계가 있습니다. CNN 모델은 여러 개체를 분류할 수 없다는 것입니다. 즉, CNN 모델은 한 번에 하나의 음식 재료만 인식할 수 있습니다. 주어진 이미지에서 다양한 음식 재료를 식별할 수 없습니다. 향후 연구에서는 여러 개의 물체를 감지할 수 있는 기능을 추가할 예정입니다.

## 참고 자료

- [1] Bergevin, R. 및 M.D. Levine, 일반 객체 인식: 선 그림에서 대략적인 설명 구축 및 일치시키기. 패턴 분석 및 기계 지능에 관한 IEEE 트랜잭션, 1993. 15(1): p. 19-36.
- [2] Bradski, G. and A. Kaehler, OpenCV. 오픈 박사의 소프트웨어 도구 저널, 2000. 3.
- [3] 아바디, M., 외. 텐서플로우: 대규모 머신 러닝을 위한 시스템. 운영 체제 설계 및 구현에 관한 {USENIX} 심포지엄({OSDI} 16). 2016.
- [4] McKinney, W., 데이터 분석을 위한 Python: Pandas, NumPy 및 IPython을 사용한 데이터 랭글링. 2012: "오라일리 미디어, Inc.".
- [5] 케트카르, N., 케라 소개, 파이썬으로 딥 러닝. 2017, Springer. p. 97-111.
- [6] Sinha, D. and M. El-Sharkawy. 씬 모바일넷: 향상된 모바일넷 아키텍처. 2019 IEEE 제10회 연례 유비쿼터스 컴퓨팅, 전자 및 모바일 통신 컨퍼런스 (UEMCON)에서. 2019. IEEE.
- [7] 타그, S., D. 알메이다, 및 K. 라이먼, 리셋에서 리셋: 잔여 아키텍처 일반화. arXiv 사전 인쇄물 arXiv:1603.08029, 2016.
- [8] 컴퓨터 비전의 시작 아키텍처에 대해 다시 생각하기. 컴퓨터 비전 및 패턴 인식에 관한 IEEE 컨퍼런스 논문집. 2016.
- [9] Qassim, H., A. Verma, and D. Feinzimer. 빅 데이터 장소 이미지 인식을 위한 압축 잔차-VGG16 CNN 모델. 2018 IEEE 8차 연례 컴퓨팅 및 통신 워크숍 및 컨퍼런스(CCWC)에서. 2018. IEEE.
- [10] Pan, S.J. and Q. Yang, 전이 학습에 관한 설문 조사. 지식 및 데이터 엔지니어링에 관한 IEEE 트랜잭션, 2009. 22(10): p. 1345-1359.
- [11] A. M. Faruk, H. A. Faraby, M. M. Azad, M. R. Fedous 및 M. K. Morol, "Deep CNN 및 양방향 게이트 반복 장치를 사용한 이미지-벡글어 캡션 생성," 2020 제 23회 국제 컴퓨터 및 정보 기술 컨퍼런스(ICCIT), 2020, 1-6쪽, doi: 10.1109/ICCIT51783.2020.9392697.
- [12] 가와노, Y. 및 K. 야나이, 푸드캠: 스마트폰의 실시간 음식 인식 시스템. 멀티미디어 도구 및 응용 프로그램, 2015. 74(14): p. 5263-5287.
- [13] Bolaños, M., M. Valdivia, and P. Radeva. 어디서 무엇을 먹는가? 이미지 기반 음식 메뉴 인식. 유럽 컴퓨터 비전 컨퍼런스. 2018. Springer.
- [14] Liu, C., 외. Deepfood: 컴퓨터 지원 식단 평가를 위한 딥러닝 기반 음식 이미지 인식. 스마트 홈 및 건강 텔레메틱스 국제 컨퍼런스. 2016. Springer.
- [15] 라보아-엑고완, D., S. 루, L. 곤잘레스, 레시피넷: 레시피/영양 정보 생성기 이미

지.

- [16] 살바도르, A., 외. 역 요리: 컴퓨터 비전 및 패턴 인식에 관한 IEEE/CVF 컨퍼런스 논문집. 2019.
- [17] hen, J., L. Pang, and C.-W. Ngo. 교차 모드 레시피 검색: 이 요리를 어떻게 요리할까요? 멀티미디어 모델링에 관한 국제 컨퍼런스에서. 2017. Springer.
- [18] 야나이, K., 마루야마 T., 카와노 Y., 식재료의 시각적 인식을 통한 요리 레시피 추천 시스템. 국제 인터랙티브 모바일 기술 저널, 2014. 8(2).
- [19] Maheshwari, S. 및 M. Chourey, 머신 러닝 모델을 사용한 레시피 추천 시스템. 국제 공학 및 기술 연구 저널 (IRJET), 2019. 6(9): p. 366-369.

- [20] 미슈라, M. 및 공, 식재료 이미지 인식을 이용한 레시피 추천 시스템. 2018.
- [21] Kingma, D.P. and J. Ba, Adam: 확률론적 최적화를 위한 방법. arXiv preprint arXiv:1412.6980, 2014.
- [22] Janocha, K. and W.M. Czarnecki, 분류에서 심층 신경망의 손실 함수에 대해. arXiv preprint arXiv:1702.05659, 2017.
- [23] 2019년 제10회 컴퓨팅, 통신 및 네트워킹 기술 국제 컨퍼런스(Iccnt)에서 정크푸드를 인식하는 머신러닝 접근법. 2019. IEEE.
- [24] Sun, F., Z. Gu, B. Feng, 이미지 특징 추출 및 분류를 통한 Yelp 음식 식별. arXiv preprint arXiv:1902.05413, 2019.
- [25] 야나이, K. 과 Y. 카와노. 사전 학습 및 미세 조정이 포함된 심층 컨볼루션 네트워크를 사용한 음식 이미지 인식. 2015 IEEE 국제 멀티미디어 및 엑스포 워크샵(ICMEW). 2015. IEEE
- [26] 박스터, J., 성분 수준 기능을 사용한 식품 인식. Ηλεκτρονικό]. Available: [http://jaybaxter.net/6869\\_food\\_project.pdf](http://jaybaxter.net/6869_food_project.pdf), 2012.
- [27] Herranz, L., S. Jiang, 및 R. Xu, 음식 인식을 위한 레스토랑 컨텍스트 모델링. IEEE 멀티미디어 트랜잭션, 2016. 19(2): p. 430-440.
- [28] Ciocca, G., P. Napoletano 및 R. Schettini, 음식 인식: 새로운 데이터 세트, 실험 및 결과. IEEE 생의학 및 건강 정보학 저널, 2016. 21(3): p. 588-598.
- [29] Meyers, A., 외. Im2Calories: 자동화된 모바일 비전 음식 일기를 향하여. IEEE 국제 컴퓨터 비전 컨퍼런스 논문집. 2015.
- [30] S. I. Sohel, C. Mondol, H. S. Ayon, U. T. Islam, M. K. Morol, "이미지 분위기 결정을 통한 음악 제안", 2021 제24회 컴퓨터 및 정보 기술 국제 컨퍼런스 (ICCIT), 2021, 1-7페이지, doi치: 10.1109/ICCIT54785.2021.9689781.