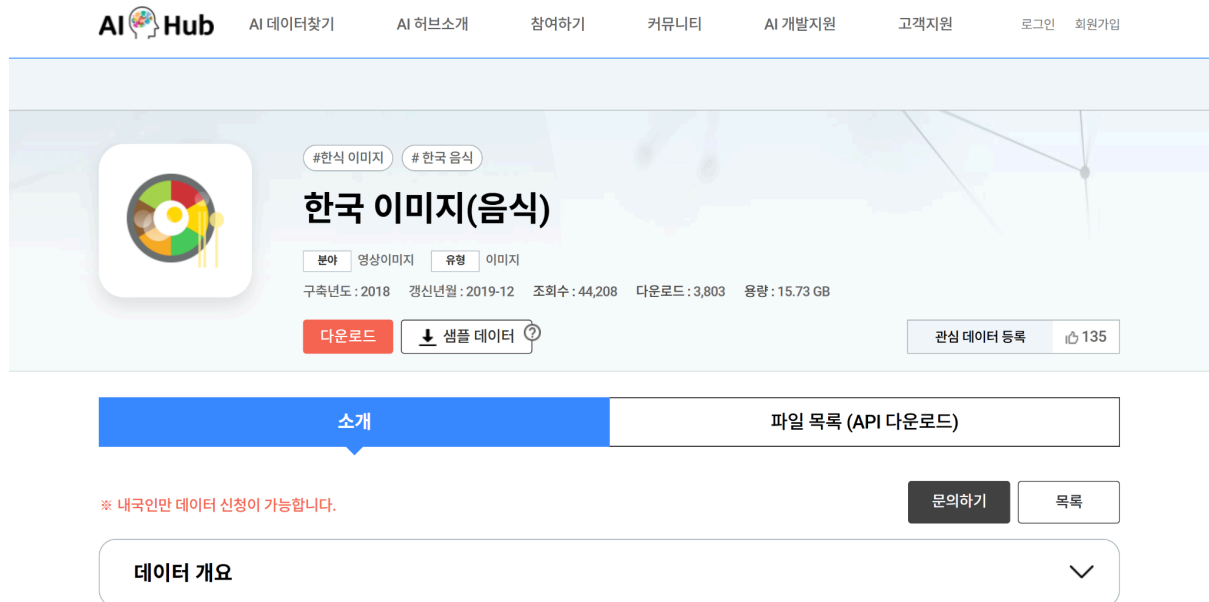


Appendix A — AI-Hub

<https://aihub.or.kr/aihubdata/data/view.do?dataSetSn=79>



AI Hub AI 데이터찾기 AI 허브소개 참여하기 커뮤니티 AI 개발지원 고객지원 로그인 회원가입

#한식 이미지 #한국 음식

한국 이미지(음식)

분야 영상이미지 유형 이미지

구축년도 : 2018 갱신년월 : 2019-12 조회수 : 44,208 다운로드 : 3,803 용량 : 15.73 GB

다운로드 샘플 데이터 ?

관심 데이터 등록 135

소개 파일 목록 (API 다운로드)

※ 내국인만 데이터 신청이 가능합니다.

문의하기 목록

데이터 개요

Appendix A1(AI-Hub dataSetSn=79)

The AI-Hub “Korean Food Images” dataset provides an image database of 150 Korean food classes, with approximately 1,000 images per class, for a total scale of 150,000 images (“150종 × 1,000장 = 150,000장”). It also states that a research-purpose usage consent form must be completed (“연구목적 활용동의서 작성 필요”) and that only a cleaned/refined subset of the constructed dataset is provided (“구축된 데이터셋 중 정제된 일부 데이터셋 제공”).

These statements justify our report’s description of AI-Hub as an access-controlled public dataset used for research/education, and support our claim that we trained our classifier using a subset of this dataset rather than redistributing the raw files.

Key phrases captured :

- “한국 음식 150종(종별 약 1천 장)” → 150 food categories, about 1,000 images per category

- “150종 × 1,000장 = 150,000장” → total dataset scale of 150,000 images
- “연구목적 활용동의서 작성 필요” → research-purpose consent form required
- “정제된 일부 데이터셋 제공” → only a refined/cleaned subset is provided

※ 내국민만 데이터 신청이 가능합니다.

문의하기

목록

데이터 개요



데이터 변경이력

버전	일자	변경내용	비고
1.0	2019-12-31	데이터 최초 개방	

소개

한국 음식 150종(종별 약 1천 장)의 데이터를 구축한 이미지 데이터 제공

- 데이터 영역 : 영상이미지
- 데이터 유형 : 이미지
- 구축년도 : 2018년
- 구축량 : 15만

데이터 통계



구축내용

- 한식 등 국내 특화 시각지능기술 및 서비스 개발을 위해 필요한 이미지 데이터의 제작, 수집, 분류 체계를 마련하고 관련 정보를 레이 블링
- 국내 인공지능 기술의 활성화를 위해서는 무엇보다 기계학습에 활용할 다양한 패턴의 고품질 이미지 데이터 확보가 중요
- 한국 음식 인식성능의 강화를 위해 한국 음식 150종(종별 약 1천장)의 데이터 수집 및 세그먼트 정보 등을 태깅

데이터 종류	구축수량	포함 내용	제공 방식
음식 이미지 데이터베이스	150종 X 1,000장 = 150,000장	각 class 별 1000장 구축	파일 다운로드

- ※연구목적 활용동의서 작성 필요
- ※구축된 데이터셋중 정제된 일부 데이터셋 제공

설명서 및 활용가이드 다운로드

데이터 설명서 다운로드

데이터 변경이력

버전	일자	변경내용	비고
1.0	2019.12.31	데이터 최초 개방	

필요성

- 해외 주요국가에서는 시각지능 분야에서 인공지능 기술개발을 위한 기계학습 데이터 구축 등 관련 인프라를 확보하여 개방
- 미국 ImageNet은 22,000개의 범주로 분류된 약 1,500만장의 이미지를 무료로 개방하여 시각 지능 기술 발전의 기초를 제공
- 국내 인공지능 기술의 활성화를 위해서는 무엇보다 기계학습에 활용할 다양한 패턴의 고품질 이미지 데이터 확보가 중요
- 기존 공개된 이미지 데이터베이스(ImageNet 등)에서 다루지 않은 국내 특화 서비스 및 데이터 개방을 위한 고유의 특성을 가진 이미지 데이터의 구축 필요

구축내용

- 한식 등 국내 특화 시각지능기술 및 서비스 개발을 위해 필요한 이미지 데이터의 제작, 수집, 분류 체계를 마련하고 관련 정보를 레이블링
- 국내 인공지능 기술의 활성화를 위해서는 무엇보다 기계학습에 활용할 다양한 패턴의 고품질 이미지 데이터 확보가 중요
- 한국 음식 인식성능의 강화를 위해 한국 음식 150종(종별 약 1천장)의 데이터 수집 및 세그먼트 정보 등을 태깅

데이터 종류	구축수량	포함 내용	제공 방식
음식 이미지 데이터베이스	150종 X 1,000장 = 150,000장	각 class 별 1000장 구축	파일 다운로드

- ※연구목적 활용동의서 작성 필요
- ※구축된 데이터셋중 정제된 일부 데이터셋 제공

데이터베이스 구조

- 한식 이미지 데이터베이스는 한식재단의 음식분류 및 한국인이 즐겨 먹는 음식통계를 참조하여 선정된 150종의 음식으로 구성되며, 한식메뉴외국어표기 길라잡이(한식재단, 200 International Korean Menu Guide, 2014)를 참고하여 음식의 종류를 대분류(밥, 면, 국 등) 및 소분류를 결정하고 ID를 부여하고 구조화

특 허 데 이 터 구 축 내 용 표 (구 축 년 도, 데 이 터 종 류, 포 함 내 용, 제	대분류	소분류	대분류	소분류
	국	갈비구이,갈치구이,고등어구이,곰창구이,닭갈비,더덕구이,떡갈비,불고기,삼겹살장어구이,조개구이,황태구이,훈제오리	국	계란국, 떡국/만두국, 무국, 미역국, 북엇국, 소고기무국, 시래기국, 육개장, 콩나물국
	김치	갓김치, 깍두기, 나박김치, 무생채, 배추김치, 백김치, 부추김치, 열무김치, 오이소박이, 총각김치, 파김치	나물	가지볶음, 고사리나물, 미역줄기볶음, 숙주나물, 시금치나물, 애호박볶음
	떡	경단	만두	만두
	면	막국수, 물냉면, 비빔냉면, 수제비, 열무국수, 잔치국수, 쫄면, 칼국수, 콩국수, 라면, 자장면, 팍팍	무침	고추된장무침, 파리고추무침, 도토리묵, 잡채, 도라지무침, 콩나물무침, 홍어무침
	밥	김밥, 김치볶음밥, 비빔밥, 새우볶음밥, 알밥, 잡곡밥, 주먹밥, 유부초밥	볶음	건새우볶음, 오징어채볶음, 감자채볶음, 고추장진미채볶음, 두부김치, 떡볶이, 라볶이, 멸치볶음, 소세지볶음, 어묵볶음, 제육볶음, 쭈꾸미볶음
	쌈	보쌈	음청류	수정과, 식혜

Appendix_A2(Terms/Policy excerpt — Portal ownership & restrictions)

This excerpt from the AI-Hub Terms of Use states that intellectual property rights related to AI-Hub services and associated materials (e.g., software, images, logos, designs, service names, information, and trademarks) belong to the operating organization (and/or contracted organizations). It also states that, **unless explicitly approved by AI-Hub**, users may not perform actions such as **modifying, renting, lending, selling, distributing, producing/creating derivative works, transferring, re-licensing, or commercially exploiting** these materials, and may not allow third parties to do so.

This supports our submission policy of **not including or redistributing raw AI-Hub content** in public repositories and limiting outputs to results (metrics/figures) and our trained model artifacts.

Key Korean lines → English meaning (for clarity):

- “지식재산권 및 기타 권리는 ... 소유권” → IP and related rights are owned by the operator/contracted 기관.
- “명시적으로 승인한 경우를 제외하고 ... 수정/대여/대출/판매/배포/양도/재라이선스/상업적 이용 불가” → Without explicit approval, various redistribution/commercial exploitation actions are prohibited.
- “제3자로 하여금 ... 허락할 수 없음” → You may not permit third parties to do the same.

이용약관

제 1조 목적	▽
제 2조 용어의 정의	▽
제 3조 약관의 효력과 변경	▽
제 4조 약관 외 준칙	▽
제 5조 이용 계약의 성립	▽
제 6조 회원 정보 사용에 대한 동의	▽
제 7조 사용자의 정보 보안	▽
제 8조 포털 이용 시간	▽
제 9조 포털의 중지 및 정보의 저장과 사용	▽
제 10조 서비스의 변경 및 해지	▽
제 11조 정보 제공 및 홍보물 게재	▽
제 12조 게시물의 저작권	▽
제 13조 사용자의 행동 규범 및 서비스의 이용 제한	▽
제 14조 포털의 의무	▽
제 15조 회원의 의무	▽
제 16조 포털의 소유권	△
<div>1. AI 허브가 제공하는 서비스, 그에 필요한 소프트웨어, 이미지, 마크, 로고, 디자인, 서비스명칭, 정보 및 상표 등과 관련된 지식재산권 및 기타 권리는 운영기관(및 AI 허브 서비스 제공과 관련하여 운영기관과 계약을 체결한 기관)에 소유권이 있습니다.</div> <div>2. 귀하는 AI 허브에서 명시적으로 승인한 경우를 제외하고는 전항의 소정의 각 재산에 대한 전부 또는 일부의 수정, 대여, 대출, 판매, 배포, 제작, 양도, 재라이선스, 담보권 설정 행위, 상업적 이용 행위를 할 수 없으며, 제3자로 하여금 이와 같은 행위를 하도록 허락할 수 없습니다.</div>	
제 17조 양도 금지	▽
제 18조 손해 조항	▽
제 19조 면책 조항	▽
제 20조 관할 법원	▽

Appendix_A3(FAQ — Commercial use)

This AI-Hub FAQ states that **AI models and services trained using data downloaded from AI-Hub may be freely distributed and used**, including for commercial purposes. However, it also requires that users **clearly disclose that AI-Hub data was used** when distributing or using such models/services. In addition, the FAQ notes that **redistributing the dataset (including reprocessed versions) is generally not allowed**, unless there is **prior consultation/approval with NIA and the dataset-building company**.

This supports our project policy: we use AI-Hub data for training and report results/metrics, but **do not redistribute raw dataset files**.

Key Korean lines → English meaning (for clarity):

- “AI모델과 서비스는 자유롭게 배포, 활용” → Models/services trained with AI-Hub data can be freely deployed and used.
- “영리적으로 판매, 활용 ... 제한하지 않으나 ... 사용하였음을 명시” → Commercial use is not restricted, but AI-Hub data usage must be disclosed.
- “데이터를 재가공하여 배포 ... 원칙적으로 불가” → Redistributing (even reprocessed) data is generally not allowed without prior agreement.

Q. AI 허브에서 제공받은 데이터를 영리적 목적으로 이용해도 되나요?



A. - AI허브에서 다운로드 받은 데이터로 학습하여 만들어진 AI모델과 서비스는 자유롭게 배포, 활용하실 수 있습니다.
AI허브 데이터를 활용해 개발한 모델을 영리적으로 판매, 활용하는 행위는 제한하지 않으나 AI허브 데이터를 사용하였음을 명시하여야 합니다.

- 다만, NIA 및 구축기업과 사전 협의한 경우를 제외하고 데이터를 재가공하여 배포하는 행위는 원칙적으로 불가합니다.