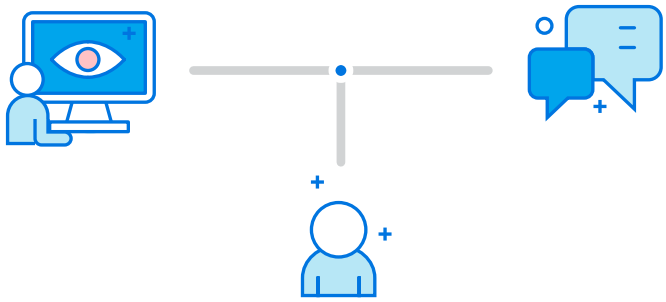
KVQA(Korean Visual Question Answering) 대규모 데이터셋

시각장애인을 위한 한국어 기반 시각적 질의응답 데이터셋

테스트웍스는 SKT 내 인공지능 연구조직 T-Brain과 함께 2019년 6월부터 12월까지 한국어 기반 시각적 질의응답 KVQA (Korean Visual Question Answering) 데이터셋을 구축했습니다.

시각적 질의응답, VQA란?

시각적 질의응답 VQA (Visual Question Answering)는 시각적인 이미지와 그 이미지에 대한 질문을 입력했을 때, 질문에 적절한 답을 제공하는 인공지능 기술입니다.



VQA 기술(출처: SK T-Brain Github, KVQA)

기존 VQA 데이터셋의 한계

VQA는 세계적으로 활발히 연구되고 있으며, 개발을 위해 나라별 언어적 특징에 맞는 이미지와 그에 대한 질문과 답변으로 구성된 학습 데이터셋이 필요한 기술입니다. 하지만 국내에는 활용할 수 있는 한국어로 된 VQA 데이터가 없었습니다. 또한, 기존 데이터셋의 대부분이 비장애인이 촬영한 이미지였습니다. 주변 환경 및 사물을 인지하기 어려운 시각장애인이 촬영한 이미지는 조도가 어둡거나 초점이 맞지 않는 등 비장애인이 촬영한 이미지와 차이가 있기 때문에 시각장애인을 위한 기술 및 서비스로 개발하는 것은 한계가 있었습니다.

세계 최초 KVQA 데이터셋

테스트웍스와 T-Brain은 이러한 VQA 기술 문제를 개선하고 시각장애인을 위한 인공지능 서비스 개발을 위해 세계 최초로 한국어 기반 VQA 데이터인 KVQA 데이터셋 구축 프로젝트를 진행했습니다.

프로젝트 기간 동안 이미지와 질문 데이터 수집에 시각장애인이 직접 참여하여 기존 VQA 데이터셋의 한계를 개선했습니다.

(a) **Q**: 지금 횡단보도를 건너도 될까?
A: 아니오

(b) **Q**: 이 방에는 몇 개의 형광들이 있나요? **A**: 2

(c) **Q**: 방에 있는 사람은 지금 뭘하고 있지? **A**: 피아노

(d) **Q**: 무슨 꽃이 피어있지?
A: 답변불가

[그림 1] KVQA 데이터셋의 예. 각 질문에 대해 가장 자주 나오는 답변이 표시됩니다. 위의 예는 예/아니오, 숫자, 기타, 답변불가 유형의 이미지 질문 쌍입니다.

KVQA 데이터셋 다운로드

KVQA 데이터(출처: SK T-Brain Github, KVQA)

시각장애인으로부터 직접 수집한 이미지 및 질문 데이터

테스트웍스는 한국시각장애인협회, 실로암시각장애인복지관, 한국시각장애대학생회, 한국시각장애인연합회 등 관련 기관들과 협력하여 이미지와 질문 데이터를 수집했습니다.

한국시각장애인연합회의 도움으로 점자 신문을 제작하는 등 적극적인 홍보 활동을 진행하여 시각장애인의 데이터 수집 참여를 독려했습니다.

데이터 수집 가공 플랫폼 aiworks를 활용한 대규모 답변 데이터 수집

테스트웍스는 데이터 수집 가공 플랫폼 aiworks를 활용해 시각장애인이 수집한 이미지와 질문데이터에 대한 답변 데이터를 수집했습니다.

먼저, KVQA 데이터 수집 프로젝트에 대한 페이지를 개설하여 답변 데이터 수집에 참여할 aiworks 회원을 모집했습니다.

KVQA 데이터 구축을 위한 랜딩페이지

그리고, aiworks에서 회원들이 이미지와 질문을 확인하고 답변을 작성하면, 답변을 작성하면 현금으로 전환가능한 포인트를 제공하거나 봉사활동 시간을 제공하는 프로젝트를 개설했습니다.

자원봉사자들을 포함한 aiworks 회원들은 PC와 스마트폰를 통해 aiworks에 접속하여 답변 데이터 수집에 참여했습니다.

aiworks를 통한 답변데이터 수집, 왼쪽 스마트폰 화면, 오른쪽 PC화면

KVQA 데이터셋 구축 성과

KVQA 데이터 구축을 통해 약 6개월 동안 전국에서 250여 명의 시각장애인이 이미지와 질문 데이터 수집에 참여했고 4,500여 명의 aiworks 회원들이 답변 데이터를 제공했습니다. 그 결과, 10만 장의 이미지, 10만 건의 질문, 그리고 100만 건의 답변에 대한 데이터를 구축할 수 있었습니다.

V1.0 (2020년 1월)

	전체 (%)	예/아니오 (%)	숫자 (%)	기타 (%)	답변불가 (%)
# images	100,445 (100)	6,124 (6.10)	9,332 (9.29)	69,069 (68.76)	15,920 (15.85)
# questions	100,445 (100)	6,124 (6.10)	9,332 (9.29)	69,069 (68.76)	15,920 (15.85)
# answers	1,004,450 (100)	61,240 (6.10)	93,320 (9.29)	690,690 (68.76)	159,200 (15.85)

KVQA 데이터셋 구축 통계(출처: SK T-Brain Github, KVQA)

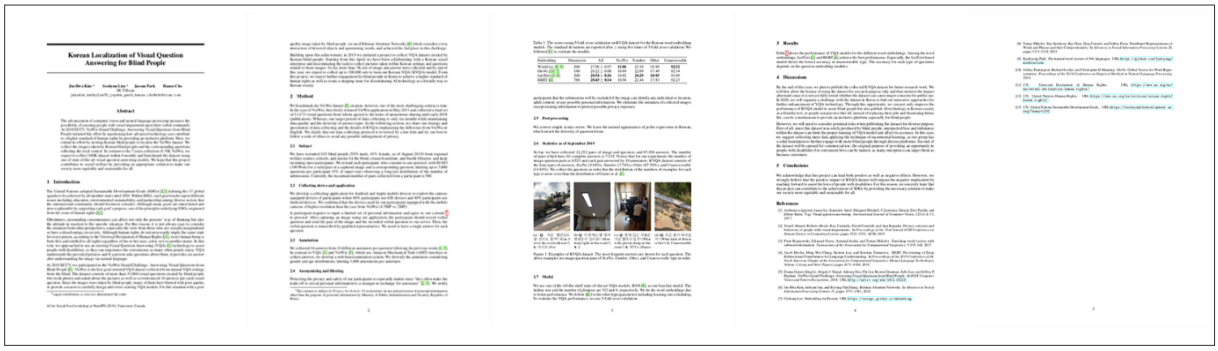
KVQA 프로젝트 이전 VQA 챌린지에서 페이스북 다음으로 2위를 차지했던 T-Brain은 구축된 KVQA 데이터셋에 대한 논문을 기고했고, 세계 최고 인공지능 컨퍼런스 2020 NeurIPS 내 워크샵 AI for Social Good에서 KVQA에 대한 성과를 소개했습니다.

테스트웍스는 KVQA 프로젝트 기간 동안 당사의 데이터 수집 가공 플랫폼 aiworks에 사회봉사 프로그램을 연계하여 사회 문제 해결을 위한 공유 가치를 창출하였고 국내 최초로 한국어 기반의 대규모 VQA 데이터셋을 성공적으로 구축했습니다.

테스트웍스는 기술 기반의 성장을 바탕으로 고객과 함께 지속적으로 사회문제 해결을 위한 다양한 공유 가치 프로젝트에 참여하고 있습니다.

논문

AI for Social Good workshop at NeurIPS (Kim & Lim et al., 2019)



KVQA 데이터셋 논문(출처: SK T-Brain Github, KVQA)

PDF 다운로드

데이터 항목

Name	Type	Description
VQA	[dict]	시각적 질의응답 정보를 담은 dict의 list
+- image	str	이미지 파일명
+- source	str	데이터 출처 ["kvqa" | "vizwiz"]
+- answers	[dict]	답변 정보를 담은 dict 10개의 list
+— answer	str	시각적 질의에 대한 답변
+— answer_confidence	str	답변에 대한 신뢰도 ["yes" | "maybe" | "no"]
+- question	str	이미지에 관련한 질의
+- answerable	int	답변 가능 여부 [0 | 1]
+- answer_type	str	답변 유형 ["number" | "yes/no" | "other" | "unanswerable"]

KVQA 데이터 항목(출처: SK T-Brain Github, KVQA)

데이터 예시


```
[{
    "image": "KVQA_190712_00143.jpg",
    "source": "kvqa",
    "answers": [{
      "answer": "피아노",
      "answer_confidence": "yes"
    }, {
      "answer": "피아노",
      "answer_confidence": "yes"
    }, {
      "answer": "피아노 치고있다",
      "answer_confidence": "maybe"
    }, {
      "answer": "unanswerable",
      "answer_confidence": "maybe"
    }, {
      "answer": "게임",
      "answer_confidence": "maybe"
    }, {
      "answer": "피아노 앞에서 무언가를 보고 있음",
      "answer_confidence": "maybe"
    }, {
      "answer": "피아노치고있어",
      "answer_confidence": "maybe"
    }, {
      "answer": "피아노치고있어요",
      "answer_confidence": "maybe"
    }, {
      "answer": "피아노 연주",
      "answer_confidence": "maybe"
    }, {
      "answer": "피아노 치기",
      "answer_confidence": "yes"
    }],
    "question": "방에 있는 사람은 지금 뭘하고 있지?",
    "answerable": 1,
    "answer_type": "other"
  },
  {
    "image": "VizWiz_train_000000008148.jpg",
    "source": "vizwiz",
    "answers": [{
      "answer": "리모컨",
      "answer_confidence": "yes"
    }, {
      "answer": "리모컨",
      "answer_confidence": "yes"
    }, {
      "answer": "리모컨",
      "answer_confidence": "yes"
    }, {
      "answer": "티비 리모컨",
      "answer_confidence": "yes"
    }, {
      "answer": "리모컨",
      "answer_confidence": "yes"
    }, {
      "answer": "리모컨",
      "answer_confidence": "yes"
    }, {
      "answer": "리모컨",
      "answer_confidence": "yes"
    }, {
      "answer": "리모컨",
      "answer_confidence": "maybe"
    }, {
      "answer": "리모컨",
      "answer_confidence": "yes"
    }, {
      "answer": "리모컨",
      "answer_confidence": "yes"
    }],
    "question": "이것은 무엇인가요?",
    "answerable": 1,
    "answer_type": "other"
  }
]
```


KVQA 데이터 예시(출처: SK T-Brain Github, KVQA)

TESTWORKS

(주)테스트웍스 www.testworks.co.kr
서울 송파구 백제고분로 41길 42-19(더블유123빌딩) 3,4층
Mail : cs@testworks.co.kr | Tel : 02-423-5178 | Fax : 02-424-5178