

테크니컬 리포트

2020년 1차
인공지능
학습용
데이터 구축

자연어 영역

수어 영상

개요: 수어(手語) 영상 데이터셋 이란?

2019년 통계청 기준 국내 청각장애인의 수는 약 37만명으로 집계된다. 대부분의 농인들은 수어를 이용하여 소통을 하고 있다. 한국 수어는 대한민국 농문화 속에서 시각 및 동작 체계를 바탕으로 생겨난, 고유한 형식의 언어이며 대한민국 농인의 공용어로 사용된다. 한국 수어는 한국어와 다른 언어이며 한국어로 이뤄진 의사소통 방식(구어, 문자)으로 농인에게 정확한 의미 전달이 어려울 수 있으며, 농인-청인 간 의사소통 및 정보전달을 위해 TV 자막, 필담 등의 대안적 방법이 주로 활용되나, 수어와 한국어의 통사 구조 차이로 인해 한계가 있다. 필담의 경우, 약 12% 농인만이 청인 수준으로 모든 의미를 완벽히 이해하고, TV 자막의 경우 약 9% 농인만이 의미를 완전히 이해가 가능한 상황이다. 특히, 농인의 문맹률은 약 30%에 달하는 것으로 추정되고 있어 한국어 문자기반의 소통이 어렵고 청인의 입장에서 농인이 작성한 글을 읽을 때 통사구조 차이로 인한 어색함을 느끼는 경우가 많으며, 수어에 대한 몰이해로 인해 소통이 어려운 경우가 많다. 이러한 소통의 한계로 인해 사회 전반에 걸친 청인 중심의 정보 전달 체계로 인해 일상생활, 경제활동, 긴급상황 등 거의 모든 분야에 걸쳐 다양한 고충을 겪으며 최종적으로 사회 통합과 참여에 있어 제한을 겪고 있으며 2008년 '장애인차별금지 및 권리구제 등에 관한 법률'(이하 장애인차별금지법)이 제정 및 시행되고 다양한 분야의 서비스 제공에 있어 차별을 금지하는 조항이 명시되어 있으나 이 조항을 지키기 위한 조치는 거의 이뤄지지 않고 있다. 특히 재해, 재난 상황 등에서 자유롭지 않은 정보 접근으로 인해 경제적 피해, 생명 피해 등 차별에 노출 빈도 및 확률이 청인에 비해 더욱 높다.

수어 영상 인식이란 농인의 수어를 인공지능 기술을 통해 의미를 정확히 이해하고 청인이 인지할 수 있는 형태(텍스트 및 오디오)로 출력하는 것을 의미하며 인공지능 기술이 앞선 미국, 중국, 유럽 국가 중심으로 Computer Vision, NLP 등을 결합 및 활용, 수어를 인식하고 농인-청인 간 의사소통의 장벽 극복을 위한 다양한 시도들이 진행 중이다. 수어 영상 데이터는 인공지능 기술을 통해 다양한 제품 및 서비스를 제공함으로써 농인들의 사회 장벽을 낮추기 위해 구축되었다.

데이터셋의 구성

본 데이터셋의 구성은 표 1. 과 같다. 최종 선정된 문장 2,000 개와 단어 3,000개, 지문자 800개, 지숫자 200개는 언어제공자를 통한 대규모 영상 데이터 수집 및 가공을 위한 견본 영상 확보를 위해 농인 연구원이 직접 촬영에 참여하여 총 6,000개의 견본 영상을 제작하여 데이터셋을 구성하였다.

제시된 표 1.의 견본 데이터셋을 통해 표 2.와 같이 한국수어 문장 2,000개와 단어 3,000개씩 20명의 언어제공자에게서 촬영장 3곳을 활용하여 총 50만개의 영상을 수집하였다.

표 1. 데이터셋 견본 구성 및 내용

수집방법	구분	한국수어 수	영상 클립 수
직접 촬영	단어	3,000	300,000 (3000 x 5각도 x 20명)
	문장	2,000	200,000 (2000 x 5각도 x 20명)
가상데이터	단어	2000	10,000 (2000 x 5각도)

Required 수어문장2000개, 단어3000개, 지문자/숫자1000개 구별자	Required 촬영방법에 따라 실촬영/가상데이터/ 클라우드소싱 수어제공자 20/1/21명 구별자	Required 촬영각도 표식자 클라우드 소싱의 경우 단방향 촬영이어서 "F"만 존재	Required Morpheme : 형태소/비수지요소 가공 XXXX_keypoints : 동영상의 해당 프레임의 키 포인트 가공
--	---	--	--

표 3.과 표 4.와 같이 수어문장/단어, 촬영방법, 촬영각도, 가공방법을 달리하여 데이터명을 구성하였다.

데이터셋의 설계 기준과 분포

한국어 자료 수집

본 데이터셋에서 한국어 자료 수집 기준은 단순 문장이 아니라 문장이 사용되는 상황을 종합적으로 파악해 의도 및 상황 목적에 맞는 문장을 선정하는 것이다. 농인들의 이동권을 보장하기 위해 서울교통공사, T-MAP 등 수어인식 인공지능의 실제 수요가 예상되는 기관 및 기업을 도메인으로 선정해 해당 기관들에서 빈도 높게 사용되는 한국어 자료를 협조 받아 한국 수어 문장 및 단어, 지문자와 지숫자 선정 과정의 기초자료로 활용할 예정이다. 한국어 담화 자료는 관련 기관 홈페이지 검색 및 협조 요청, 문헌 조사 등의 방법으로 수집하였다.

수어 통역사 한국어 번역

해당 도메인에서 빈도 높게 사용되어 수집된 한국어 자료를 기반으로 한국어와 한국수어에 능통한 수어 통역사가 농인과 청인의 담화를 예상하여 한국어 문장 2,000개를 선정하였다. 선정한 한국어 문장 2,000개의 의미를 분석하여 번역하였으며, 각각 한국수어 영상으로 녹화하여 기록하였다. 또한 한국어로 표현된 문장의 이해에 있어서 농인과 청인 간 차이가 존재하며 선정한 한국어 문장이 가지는 정확한 의미 전달을 위해 농인의 참여에 앞서 한국어와 한국수어에 능통한 수어 통역사의 번역과정이 진행되었다.

한국 수어 문장 선정 및 농인 교차 검토

수어를 제1언어로 하는 농인 연구원은 (나)에서 촬영된 한국수어 영상 자료를 보고 현장에서 실제 사용되는 한국수어 어휘와 문형을 고려하여 수정 및 확정하며, 수정된 내용은 영상으로 촬영하였다. 한국수어도 독립적 언어이므로 같은 의미가 발화자마다 어휘나 문형이 조금씩 달라 다양한 문장으로 표현될 수 있다.

예) 한국어: 서울역으로 가주세요.

한국수어1: [서울] [기차①] [곳]

한국수어2: [서울] [기차②] [곳]

한국수어3: [서울] [KTX] [곳]

한국수어4: [서울] [오래 ㄱ①]

한국수어5: [서울] [ㅇ ㄷ ㄱ ②]

한국수어6: [기차①/②] [서울] [곳/ㅇ ㄷ ㄱ ①/②]

한국수어7: [KTX] [서울] [곳/ㅇ ㄷ ㄱ ①/②]

한국수어 단어 선정 및 수어 번역

한국 수어 문장 2,000개에서 단어를 추출하며, 변이형과 일치동사, 분류사도 포함하여 선정하였다.

예) 한국어: 서울역으로 가주세요.

한국수어1: [서울] [기차①] [곳]

한국수어2: [서울] [기차②] [곳]

한국수어3: [서울] [KTX] [곳]



한국수어4: [서울] [ㅇ ㄷ ㄱ ①]

한국수어5: [서울] [ㅇ ㄷ ㄱ ②]

한국수어6: [기차①/②] [서울] [곳/ㅇ ㄷ ㄱ ①/②]

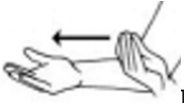
한국수어7: [KTX] [서울] [곳/ㅇ ㄷ ㄱ ①/②]

또한 한국수어의 변이형과 이형동의어, 분류사 등 표현 중 사용 빈도가 높은 것들을 중심으로 문장에 적용하여 추가 추출하였다.


변이의 예) 기차: [], []

ㅇ: [], []


이형동의어의 예) 오만원 [], []


분류사의 예) 주다 기본형: []


(돈을) 주다: []


(가방을) 주다: []


도메인과 관련한 그 밖의 한국수어 단어와 지문자, 지숫자도 추가 추출하며, 추출한 단어에서 동형이의어와 이형동의어, 합성어와 파생어를 구분하여 표기하였다.


동형이의어의 예) 주민센터와 미인대회 []


일치동사의 예) (당신이 나에게) 말한다: []


(내가 당신에게) 말한다: []

단일어2: 곳 []

합성어: 부산역 []

파생어의 예) 독립형태소: 3 []

의존형태소: 寸 []

파생어: 삼촌 []

모든 한국수어 단어는 영상과 문서로 기록하며, 농인 연구원의 교차 검토를 통해 단어 3,000개와 지문자 800개, 지숫자 200개를 최종 선정하였다. 지명과 지하철역명 등과 같은 고유명사는 한국 수어로 모두 표현되지 않고 주로 지문자로 표현되고 있다. 또한 일상생활에서 날짜, 시간, 금액, 개수 등을 표현하기 위해 지숫자는 빈번하게 사용되는 수어이다. 이에 지문자 800개와 지숫자 200개를 본 과제의 수어 단어 범위에 포함하였다. 택시와 지하철 상황을 중심으로 빈도 높게 사용될 수 있는 지숫자 200개와 지문자 800개를 추출하여 영상으로 기록하였으며, 획득한 문장 2,000개와 단어 3,000개, 지숫자 200개, 지문자 800개는 농인 연구진이 교차 검토하여 최종 확정하고, 총 6,000개의 샘플 영상을 제작하는 것으로 정제하였다.

인공지능 학습용 데이터 설계

멀티카메라 시스템 및 보정을 통한 폐색(Occlusion)에 강건한 데이터셋 구축과 수어의 통사 구조 및 비수지 특징을 반영한 데이터셋 구축으로 품질 향상을 데이터 설계 목표로 설정하였다.

(1) 수어 영상 내 특징점 추출 데이터

영상만을 사용하여 언어제공자의 수동, 수형, 수향, 수위 및 표정을 파악하기 위해 수어 동영상의 각 프레임 별 상반신(손, 팔, 얼굴)의 2차원 특징점(Keypoints) 추출한다.

다중 심점 영상으로부터 3차원 특징점을 추출하고 이를 다시 각 시점 영상에 투영해 폐색에 강건한 2차원 특징점 좌표를 취득함으로써 언어제공자의 수어 발화 과정에서 손이 얼굴 또는 반대 손을 가리며 발생하는 자기 폐색 문제(Self-occlusion)을 완화한다.

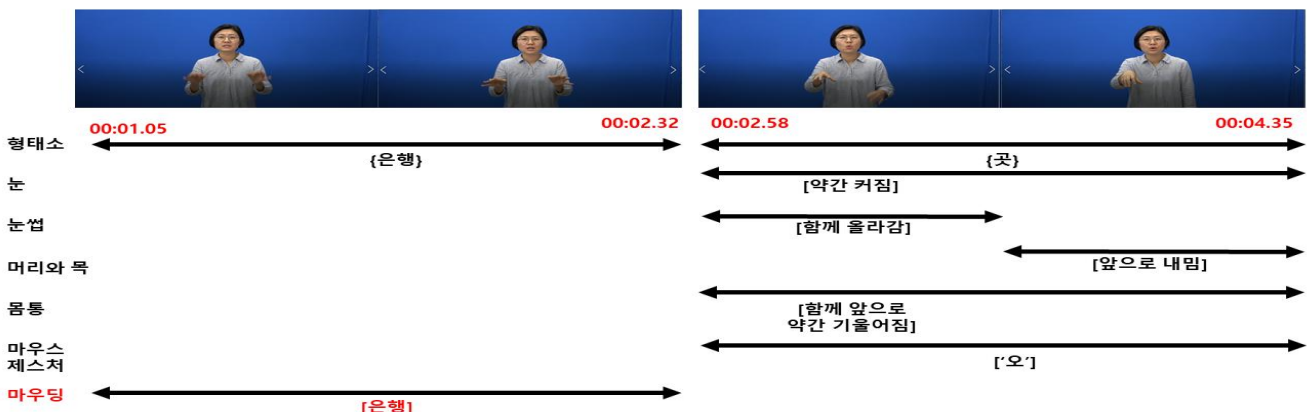
(2) 수어 형태소 단위 가공 데이터

수어-한국어 번역을 위한 언어 모델 학습에 한 개 이상의 수어 형태소로 이루어진 수어 단어/문장과 그에 대응하는 한국어 단어/문장으로 구성된 병렬 말뭉치(Parallel corpus) 데이터가 필요하다.

수어 형태소 단위 언어 모델을 구축하기 위하여 구축된 수어 단어/문장 동여상을 형태소 단위로 반할하고, 수어 형태소 간 전이관계(Transition) 추출 필요하다. 형태소 뿐 아니라 문장 유형을 구분하는데 활용되는 비수지 수어 및 단어 수준에서 표현되는 비수지 수어에 대한 annotation도 함께 수행한다.

사진 2 말뭉치 영상에 대한 설명

“은행이 어디에 있나요?”



데이터 구조

수어 영상 AI 데이터 구조

데이터셋은 {수어 단어/문장 동영상, 특징점 추출 json파일(openpose), 형태소/비수지 정보를 담은 json 파일}로 구성한다.

사진 3. 수어 데이터 구조 예시



사진 4. 언어제공자의 2차원 자세 정보

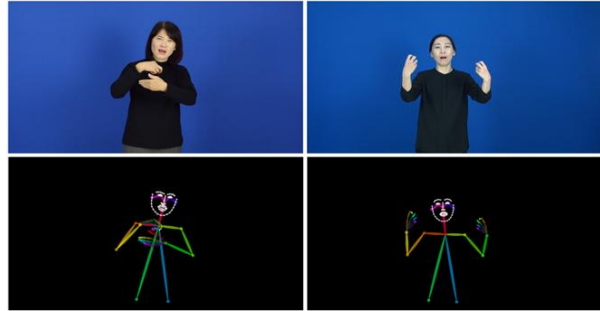


표 5. 데이터 파일 형식

형태소/비수지 json 파일	특징점 추출 json 파일(openpose)
<pre>{ "0": { "metaData": { "url": "http://52.231.159.158:1337/uploads/75373b840a304c149e88467af4344d12.mp4", "exportedOn": "2020/06/10" }, "data": [{ "start": "3.195", "end": "4.857", "attributes": [{ "name": "인사", "attribute": ["두 손 엄지 위로"] }] }] } }</pre>	<pre>{"version": 1.2, "people": [{"pose_keypoints_3d": [], "hand_left_keypoints_2d": [1103.22, 999.152, 0.570302, 1076.33, 1021.23, 0.700952, 1060.01, 1056.75, 0.786553, 1054.25, 1095.15, 0.596289, 1048.49, 1123, 0.347944, 1085.93, 1082.67, 0.656051, 1080.17, 1124.92, 0.521479, 1076.33, 1148.92, 0.436131, 1073.45, 1164.28, 0.202921, 1106.1, 1087.47, 0.568555, 1101.3, 1124.92, 0.435705, 1096.49, 1153.72, 0.240531, 1093.61, 1170.04, 0.180727, 1116.66, 1082.67, 0.574472, 1117.62, 1116.27, 0.413878, 1121.46, 1136.44, 0.226742, 1115.7, 1150.84, 0.119373, 1123.38, 1075.95, 0.459594, 1130.1, 1101.87, 0.365101, 1132.02, 1120.12, 0.280733, 1132.02, 1132.6, 0.204882], "face_keypoints_2d": [811.265, 304.002, 0.824963, 812.075, 325.054, 0.820251, 813.694, 344.485, 0.848042, 815.313, 364.727, 0.898498, 820.171, 384.968, 0.837134, 827.458, 402.781, 0.852965, 840.413, 417.355, 0.900078, 854.987, 430.309, 0.843596, 876.848, 435.977, 0.764784, 899.518, 432.738, 0.794696, 919.759, 424.642, 0.823367, 9</pre>

<pre>] }, { "start": "5.845", "end": "8.574", "attributes": [{ "name": "안녕하세요" }] }, { "start": "10.732", "end": "12.444", "attributes": [{ "name": "만나다", "attribute": ["두손 검지 모아서"] }] }, { "start": "15.337", "end": "18.574", "attributes": [{ "name": "만나서" }] } </pre>	<pre> 38.382, 411.687, 0.868127, 950.526, 397.113, 0.830881, 9 57.813, 377.681, 0.843274, 963.481, 358.25, 0.815254, 9 65.1, 336.389, 0.794279, 969.149, 315.338, 0.7923, 828.2 68, 287, 0.881878, 839.603, 280.522, 0.844393, 852.558, 277.284, 0.873537, 865.512, 280.522, 0.87778, 876.848, 2 84.571, 0.808952, 900.328, 287, 0.851515, 913.282, 281. 332, 0.929541, 926.237, 281.332, 0.904316, 938.382, 28 4.571, 0.848882, 948.097, 294.286, 0.807395, 885.754, 3 07.241, 0.855895, 884.944, 319.386, 0.832148, 884.134, 3 29.102, 0.808684, 883.325, 340.437, 0.848889, 869.561, 3 53.392, 0.874303, 876.038, 353.392, 0.865505, 881.705, 3 55.821, 0.830829, 889.802, 354.201, 0.849064, 897.089, 3 53.392, 0.866624, 840.413, 307.241, 0.954701, 847.7, 30 2.383, 0.923747, 859.035, 304.812, 0.841254, 866.322, 3 12.909, 0.897021, 857.416, 312.909, 0.856102, 846.89, 3 11.289, 0.887033, 908.424, 313.718, 0.900816, 917.33, 3 08.051, 0.942402, 925.427, 308.051, 0.996126, 932.714, 3 13.718, 0.957743, 925.427, 314.528, 0.988759, 917.33, 3 14.528, 0.951866, 857.416, 380.92, 0.83081, 866.322, 37 2.014, 0.964918, 873.609, 366.346, 0.935155, 880.896, 3 68.775, 0.863327, 888.183, 366.346, 0.867219, 898.708, 3 72.823, 0.916311, 905.995, 384.968, 0.932316, 898.708, 3 91.446, 0.92859, 887.373, 392.255, 0.92234, 880.086, 39 2.255, 0.942877, 872.799, 392.255, 0.971283, 864.703, 3 89.017, 0.816575, 861.464, 380.11, 0.880355, 873.609, 3 78.491, 0.936962, 880.086, 379.301, 0.94706, 887.373, 3 79.301, 0.910977, 901.137, 384.159, 0.831621, 887.373, 3 79.301, 0.879087, 880.086, 379.301, 0.940059, 873.609, 3 79.301, 0.946586, 853.367, 307.241, 0.902516, 919.759, 3 10.48, 0.865461], "hand_right_keypoints_3d": [], "face_ke ypoints_3d": [], "pose_keypoints_2d": [883.47, 338.986, 0. </pre>
---	---

"name": "반가워요"	94006, 889.32, 518.526, 0.737245, 733.343, 509.729, 0.6
}	79976, 633.266, 689.145, 0.752801, 674.458, 439.013, 0.
]	801585, 1051.29, 527.351, 0.671717, 1095.32, 762.803, 0.
}	603583, 1101.26, 998.255, 0.785585, 871.742, 1048.31, 0.
]	287957, 762.834, 1045.38, 0.221621, 0, 0, 0, 0, 0, 0, 977.
}	611, 1051.26, 0.231996, 0, 0, 0, 0, 0, 0, 856.908, 306.646,
}	0.876022, 924.559, 312.477, 0.871423, 809.88, 324.258,
	0.900256, 974.676, 338.888, 0.942325, 0, 0, 0, 0, 0, 0,
	0, 0, 0, 0, 0, 0, 0, 0, 0, 0], "hand_right_keypoints_2d":
	[677.062, 441.396, 0.751231, 653.069, 423.718, 0.84315
	4, 632.865, 406.039, 0.738844, 620.237, 392.148, 0.8510
	64, 602.558, 383.309, 0.917243, 656.857, 363.105, 0.768
	874, 661.909, 329.01, 0.824041, 663.171, 303.754, 0.796
	301, 666.96, 283.55, 0.77469, 682.113, 361.842, 0.93918
	3, 683.376, 321.433, 0.824474, 684.638, 293.652, 0.8216
	18, 687.164, 270.922, 0.937464, 702.317, 365.63, 0.8455
	83, 703.58, 329.01, 0.932438, 703.58, 301.229, 0.920839,
	703.58, 279.762, 0.927401, 719.996, 376.995, 0.784026, 7
	23.784, 346.689, 0.846252, 730.098, 330.272, 0.865189, 7
	32.624, 311.331, 0.95114], "hand_left_keypoints_3d": []]}

표 6. json 포맷 요소

항목	설명
metaData	원천 영상 파일명
start	수어 문장/단어의 각 형태소 시작 시각
end	수어 문장/단어의 각 형태소 종료 시각
attributes(name)	형태소 정보
attributes(attributes)	비수지 정보

hand_left_keypoints_2d	왼손 특징점 좌표 값
hand_right_keypoints_2d	오른손 특징점 좌표 값
face_keypoints_2d	얼굴 내 특징점 좌표 값
pose_keypoints_2d	몸통, 팔 등 자세 특징점 좌표 값

사진 5. 특징점(Keypoint), Annotation 파일(.json)

```

{"version": "1.3", "people": [{"person_id": 1,
  "pose_keypoints_2d": [936.474, 309.586, 0.840106, 933.506, 489.141, 0.751833, 774.534, 494.932, 0.652396, 727.5, 759.824, 0.681051, 733.385, 971.781, 0.773203, 1095.5, 494.929, 0.676582,
  1145.47, 751.064, 0.636569, 1142.57, 974.633, 0.82767, 933.566, 1018.83, 0.39798, 836.356, 1015.92, 0.366143, 0.0, 0.0, 0.0, 1042.44, 1024.73, 0.347738, 0.0, 0.0, 0.0, 906.931, 274.241, 0.928095,
  974.696, 274.177, 0.916439, 859.81, 303.615, 0.884054, 1024.66, 297.72, 0.940167, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0],
  "face_keypoints_2d": [858.516, 281.907, 0.914281, 858.516, 302.114, 0.907168, 860.133, 321.513, 0.851506, 864.174, 340.912, 0.912918, 871.449, 360.312, 0.914759, 884.382, 375.669, 0.824924,
  902.972, 386.177, 0.877267, 920.755, 393.452, 0.842267, 941.77, 396.685, 0.869358, 961.169, 392.643, 0.873439, 980.568, 385.369, 0.872462, 995.118, 373.244, 0.86357, 1007.24, 358.695, 0.872073,
  1012.9, 340.104, 0.899476, 1015.33, 320.705, 0.863703, 1016.13, 301.306, 0.856774, 1015.33, 281.099, 0.838721, 877.915, 250.384, 0.889752, 888.423, 244.726, 0.837206, 898.123, 245.534, 0.835413,
  910.247, 249.575, 0.90827, 922.371, 252.809, 0.83112, 952.278, 254.425, 0.794669, 962.786, 248.767, 0.861648, 974.91, 244.726, 0.906609, 987.843, 243.917, 0.908557, 999.159, 250.384, 0.856934,
  936.921, 275.441, 0.872288, 938.537, 288.373, 0.890619, 940.154, 301.306, 0.886794, 940.962, 314.239, 0.915199, 928.029, 326.363, 0.898359, 934.496, 327.171, 0.964298, 940.962, 327.98, 0.956743,
  947.428, 327.171, 0.918602, 953.086, 324.747, 0.819395, 890.848, 277.057, 0.919349, 898.123, 274.632, 0.891444, 909.439, 275.441, 0.932059, 916.713, 281.099, 0.887848, 909.439, 281.907, 0.973413,
  898.123, 281.907, 0.938061, 956.32, 279.482, 0.880896, 964.403, 273.824, 0.863955, 974.102, 273.016, 0.919652, 981.377, 276.249, 0.953962, 974.91, 281.099, 0.949198, 966.019, 281.907, 0.901901,
  915.097, 348.187, 0.87314, 923.18, 344.146, 0.901337, 932.879, 340.104, 0.880583, 941.77, 340.912, 0.933012, 948.237, 340.104, 0.93458, 956.32, 340.912, 0.895492, 966.827, 347.379, 0.925488,
  957.128, 353.037, 0.861229, 949.045, 353.845, 0.92849, 941.77, 355.462, 0.897918, 932.879, 355.462, 0.858779, 923.18, 353.845, 0.965702, 919.946, 348.187, 0.810963, 932.879, 347.379, 0.880667,
  941.77, 347.379, 0.947931, 949.045, 346.571, 0.953535, 961.169, 347.379, 0.909085, 949.045, 346.571, 0.967149, 941.77, 347.379, 0.950677, 932.879, 347.379, 0.894965, 903.781, 276.249, 0.928394, 968.444, 275.441, 1.02544],
  "hand_left_keypoints_2d": [1143.52, 971.823, 0.610985,
  1115.76, 990.97, 0.756975,
  1097.57, 1026.39, 0.8112,
  1093.74, 1059.9, 0.856497,
  1096.61, 1084.79, 0.798886,
  1123.42, 1052.24, 0.853489,
  1120.55, 1090.53, 0.775881,
  1111.93, 1107.77, 0.53473,
  1106.19, 1117.34, 0.358118,
  1145.44, 1054.15, 0.830878,
  1138.74, 1091.49, 0.805042,
  1128.21, 1105.85, 0.40135,
  1121.51, 1109.68, 0.390946,
  1159.8, 1051.28, 0.767793,
  1154.06, 1083.83, 0.666356,
  1145.44, 1096.28, 0.53453,
  1133.95, 1097.23, 0.331486,
  1168.42, 1042.67, 0.795916,
  1165.54, 1073.3, 0.789253,
  1160.76, 1083.83, 0.651792,
  1153.1, 1090.53, 0.481594],
  "hand_right_keypoints_2d": [730.393, 967.706, 0.662761, 755.067, 996.328,
  747.171, 1095.02, 0.624747, 760.988, 1112.79, 0.475612, 769.871, 1118.71, 0.
  714.602, 1053.57, 0.729009, 722.497, 1083.18, 0.559338, 738.288, 1093.05, 0.578976, 749.145, 1094.04, 0.327996, 709.667, 1044.69, 0.676138, 718.549, 1073.31, 0.620667, 730.393, 1079.23, 0.521587, 738.288, 1084.17, 0.366937],
  "pose_keypoints_3d": [],
  "face_keypoints_3d": [],
  "hand_left_keypoints_3d": [],
  "hand_right_keypoints_3d": []}]
}

```

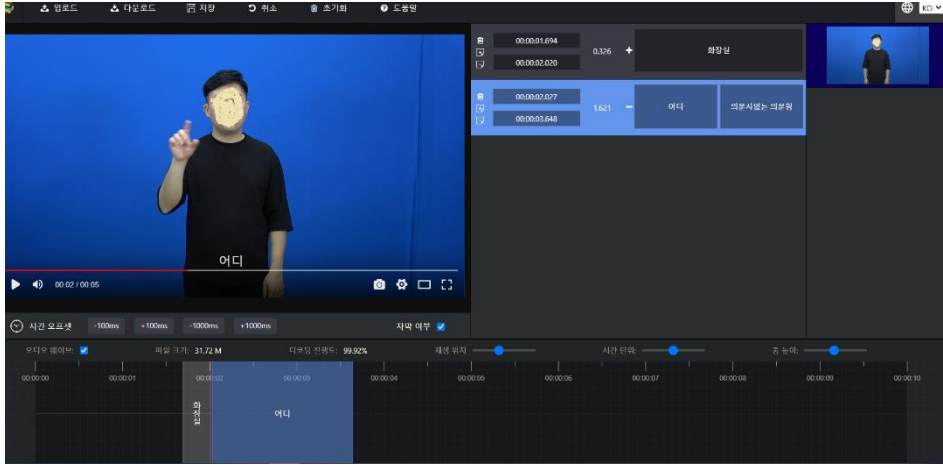
Annotation 정보

왼손 21개 키포인트 각각의 x좌표값, y좌표값, confidence값

오른손 21개 키포인트 각각의 x좌표값, y좌표값, confidence값

사진 6. 형태소/비수지 Annotation(.json)

```
{
  "0": {
    "metaData": {
      "url": "http://52.231.159.158:1337/uploads/b6077a5fd34e409593f7ee71ed22be5d.MP4",
      "duration": 5.005,
      "exportedOn": "2020/07/15",
      "data": [
        {
          "start": 1.694,
          "end": 2.020,
          "attributes": {
            "name": "화장실"
          }
        },
        {
          "start": 2.027,
          "end": 3.648,
          "attributes": {
            "name": "어디",
            "attribute": ["의문사없는 의문형"]
          }
        }
      ]
    }
  }
}
```



위 사진 6. 과 같이 수어의 언어 구성 요소에 얼굴 표정이 들어가므로 수어 인식을 위해서는 표정 변화를 확인할 수 없는 얼굴 비식별화 과정을 거칠 수 없지만, 이 자료에는 개인정보 보호를 위해 비식별화 처리하였다. 그리고 다양한 비수지 요소에 대해 표준안이 마련되어 있지 않은 상황이고, 비수지 요소가 수어영상 데이터셋에서 고려된 사례도 세계적으로 전무한 상황 아래 공통성을 발견할 수 있는 “의문사가 없는 의문형”에 대한 비수지 가공을 우선적으로 수행한다.

위 표 6. 의 상세 정보를 담은 표

표 7. 공통정보 <metaData>

Tag	설명
url	수어동영상 경로
duration	수어동영상 전체 구간
exportedOn	json 다운로드 받은 날짜

표 8. 어노테이션 정보 <data>

Tag	Attribute	설명
-----	-----------	----

start	-	해당 형태소 구간의 시작 시각
end	-	해당 형태소 구간의 종료 시각
attributes	name	해당 형태소의 표제어
	attribute	해당 형태소의 비수지 요소

데이터 예시

사진 7. 가공 완료된 데이터 xml 추출

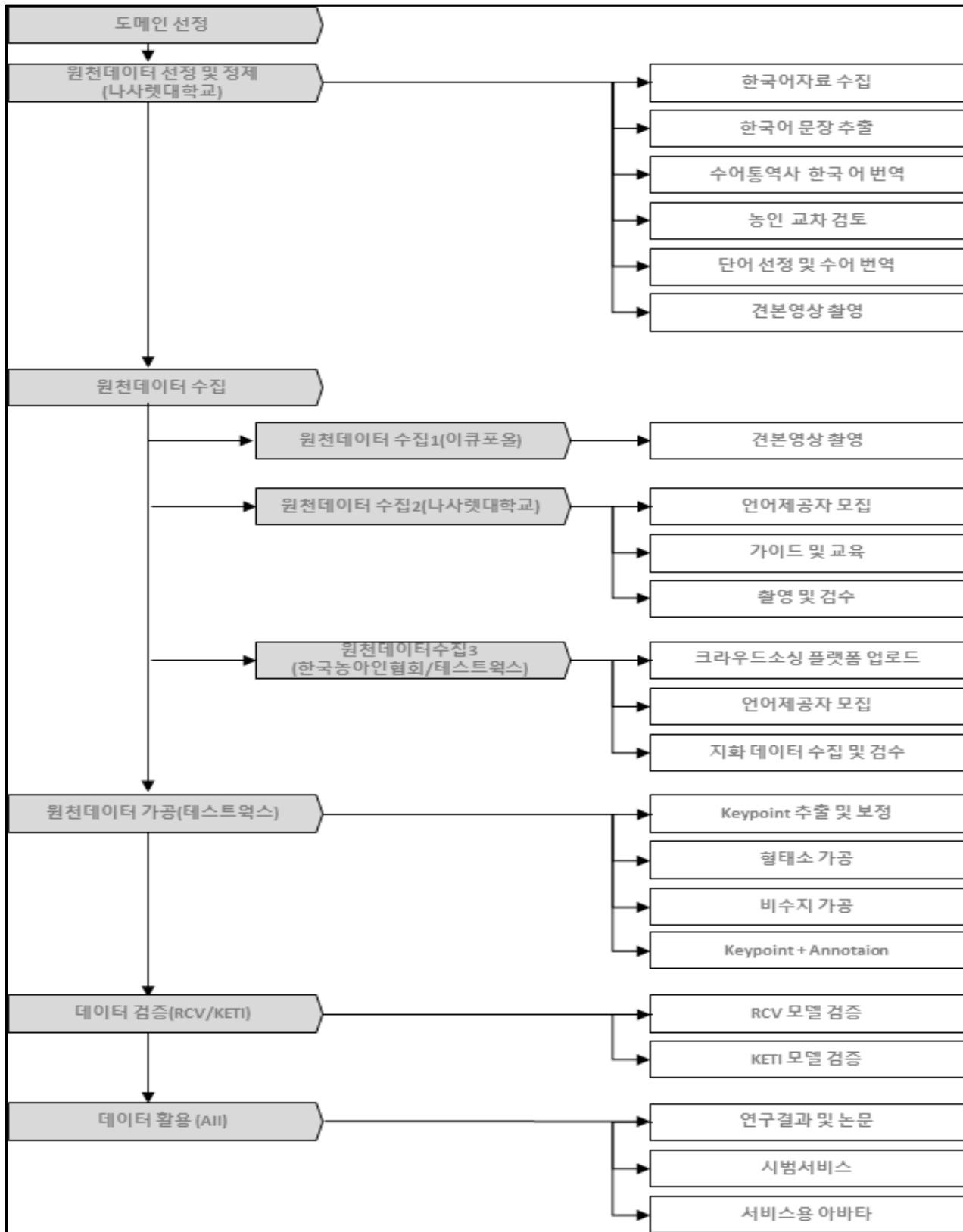
```

<?xml version="1.0" encoding="UTF-8"?>
<annotations>
  <version>1.1</version>
  - <meta>
    - <task>
      <id>16233</id>
      <project>NIA_SL</project>
      <name>NIA_SL_WORD0321_REAL09_D.mp4</name>
      <size>129</size>
      <mode>interpolation</mode>
      <overlap>0</overlap>
      <bugtracker/>
      <flipped>False</flipped>
      <created>2020-10-06 15:14:38.861408+09:00</created>
      <updated>2020-10-14 15:59:39.281097+09:00</updated>
      <start_frame>0</start_frame>
      <stop_frame>128</stop_frame>
      <frame_filter>step=1</frame_filter>
    + <labels>
    - <segments>
      - <segment>
        <id>16231</id>
        <start>0</start>
        <stop>128</stop>
        <url>http://20.41.97.88:8080/?id=16231</url>
      </segment>
    </segments>
    - <owner>
      <username>testworksadmin</username>
      <email>bo@testworks.co.kr</email>
    </owner>
    - <original_size>
      <width>2048</width>
      <height>1152</height>
    </original_size>
  </task>
  <dumped>2020-10-21 12:08:01.029024+09:00</dumped>
  <source>/home/django/data/16233/upload/NIA_SL_WORD0321_REAL09_D.mp4</source>
</meta>
- <track group_id="1" label="hand_left_keypoints_2d" id="0">
  <leftHand z_order="3"
  points="1128.570,1095.460,0.58134;1101.700,1111.580,0.62370;1084.510,1152.000,0.59
  keyframe="1" occluded="0" outside="0" frame="0"> </leftHand>
  <leftHand z_order="3"
  points="1128.570,1095.460,0.58134;1101.700,1111.580,0.62370;1084.510,1152.000,0.59

```

데이터 구축 과정

사진 8. 수어 영상 AI 데이터셋 구축 및 활용 공정



데이터 구축은 협약일 2020년 5월 7일부터 12월 31일(7개월)까지 최종 선정된 문장 2,000개와 단어 3,000개, 지문자 800개, 지숫자 200개를 언어제공자를 통한 대규모 영상 데이터 수집 및 가공을 위한 견본 영상 확보를 위해 농인 연구원이 직접 촬영에 참여하여 총 6,000개의 견본 영상을 제작하여 데이터셋을 구성을 하며 위 상기 제시된 표 1.의 견본 데이터셋을 통해 표 2.와 같이 한국수어 문장

2,000개와 단어 3,000개씩 20명의 언어제공자에게서 촬영장 3곳을 활용하여 총 52만 1천개의 영상을 수집하여 데이터셋을 구축하였다. 그리고 3가지 수어 영상 데이터 수집 방안을 통한 향후 연구 및 향후 활용도 제고와 자동화 도구를 통한 데이터셋 구축 효율성과 데이터셋 가공 전문인력을 통한 고품질 데이터셋 구축에 최우선 목표로 삼았다.

사진 8. 수어영상 데이터 수집 및 가공 절차

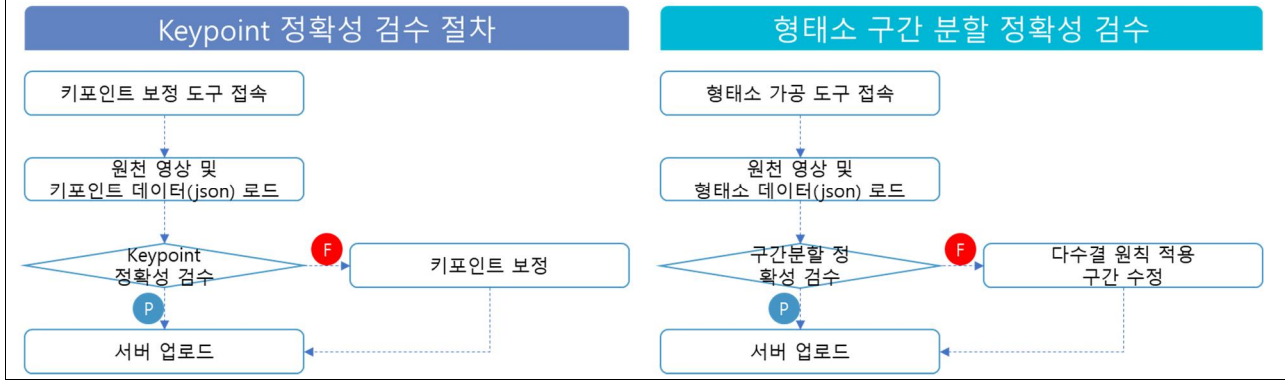


검수와 품질 확보

본 과제에서 데이터 가공시 필요한 검수는 1) 특징점(Keypoint) 추출의 수정 값 정확성 검수와 2) 형태소 구간 분할의 정확성 검수가 있다. 각 검수 절차 및 검수 기준 미달 데이터 처리 절차는 아래 표와 같다.

순번	절차	내용
1	특징점(Keypoint) 정확성 검수	Pose Estimation 도구의 경우, 성능 개발이 현재 진행 중이기에 완벽한 특징점(Keypoint) Detection 이 현실적으로 되기 어려움 이에, 자체 개발한 도구를 활용해 Frame 단위로 영상 및 Keypoint 좌표를 육안으로 검수해 Keypoint가 신체 부위에 정확하게 추출되지 않은(미추출, 오추출) 이미지 검출
2	특징점(Keypoint) 보정	자체 개발한 Keypoint 보정 도구를 활용해, 신체 해당 지점에 맞도록 좌표 추가 혹은 수정 작업 수행
3	구간 분할 정확성 검수	수어 동작이 연속적으로 진행되는 경우, Transition 발생 등, 특정 의미를 가지는 수어의 시작과 끝을 정확히 구분하기가 어려움 데이터셋의 품질 확보를 위해 분할된 구간이 정확한 지 검수 수행

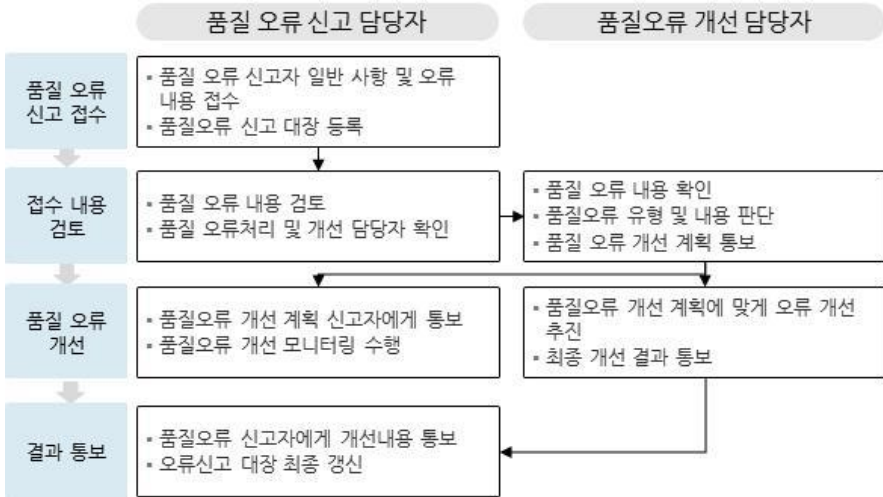
4	다수결 원칙 적용, 구간 수정	수어를 제1언어로 사용하는 농인 어노테이터 3인이 구간을 동시에 확인하고 다수결 원칙에 따라 구간 수정
5	서버 업로드	보정된 Keypoint 좌표 및 구간 정보는 학습용 데이터로 저장되어 서버 업로드



검수 기준은 아래 표와 같다.

검수절차	통과 조건	중점 확인 내용	구분
키포인트 정확성	주요 관절 지점에 키포인트의 정확한 추출	누락 포인트 및 인체 구조상 불가능한 포인트가 추출될 시, 좌표 스크립트 분석을 통해 검수 대상 이미지 확인	자동
		모든 키포인트가 추출되어 있는지 확인(미추출) 키포인트가 정확한 위치에 추출되어 있는지 확인(오추출)	육안
구간분할 정확성	형태소 발화 구간의 정확한 구분	분할 구간에 대해 3인의 어노테이터의 의견이 일치하는 지 확인	육안
형태소 정확성	정확한 형태소 입력	수어 형태소 표제어 Glossary 대로 가공이 되었는지 확인	육안

데이터 품질 오류에 대한 개선 사항은 아래의 표와 같다.



데이터 구축 담당자

수행기관(주관) : (주)테스트웍스 (전화: 02.423.5178), 이메일: ex)AI@testworks.co.kr