

진경민 직무역량기술서



고려대학교 인공지능학과

2021년 10월 27일

Experience

▪ Frontend Developer

- Deer Corporation – 전동 키펀드 스타트업
- KLUE - 고려대 강의평가 서비스

▪ AI Research

- 포즈 추정 연구 확장 – 2022.05 ~
- Sparsely-Labeled Videos 포즈 추정 IEEE SMC (Oral) – 2022.10 예정
- 미소 AI Challenge 2021 대상 – 2021.12
- 보이스 캐디 골프 스윙 분석 – 2021.05 – 2021.10
- 인공지능학과 학석사 연계과정 입학 – 2021.02
- 인공지능 융합전공 – 2019.09 – 2021.06

■ 연구 제목

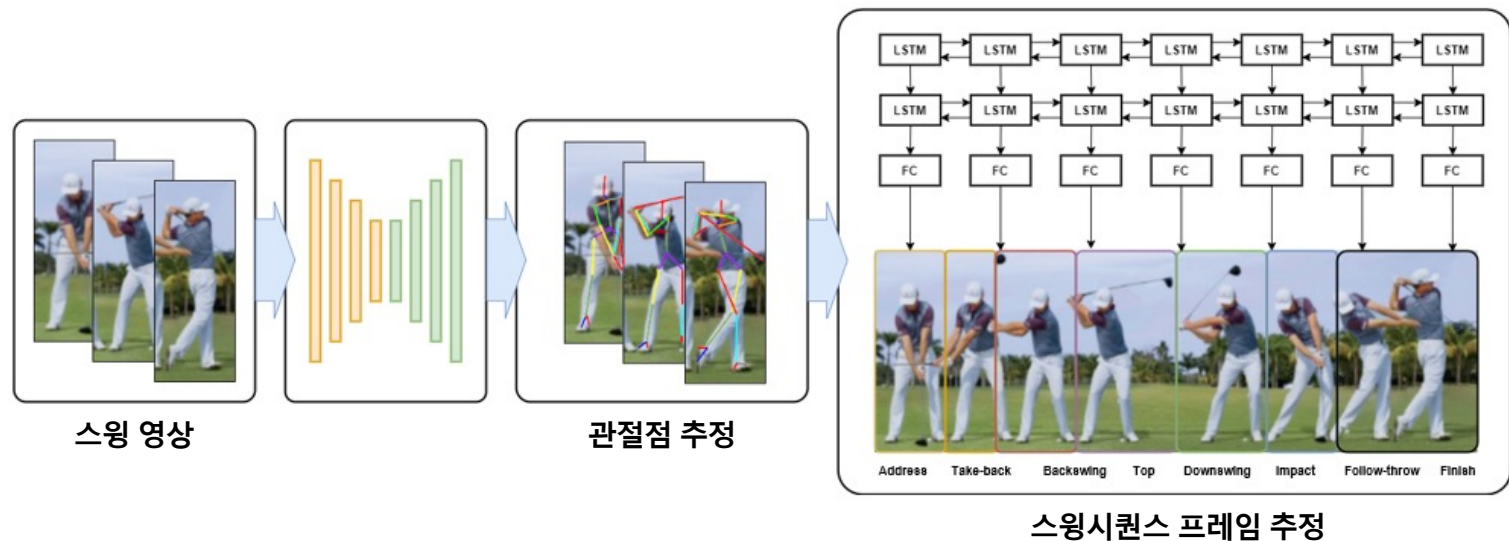
- 골프 트레이닝을 위한 인공지능 기반 골프 스윙 분석 알고리즘 개발

■ 연구 기간

- 2021. 05. 01 ~ 2021. 10. 31 (6개월)

■ 연구 목표

- 골퍼의 스윙 영상에서 관절점 추정 알고리즘 개발
- 골퍼의 스윙 영상에서 주요 스윙시퀀스 프레임 추정 알고리즘 개발
- 골퍼의 스윙 영상에서 Annotation 작업 수행



주제 1: 스윙 영상에서 관절점 추정 알고리즘 개발

■ 목표

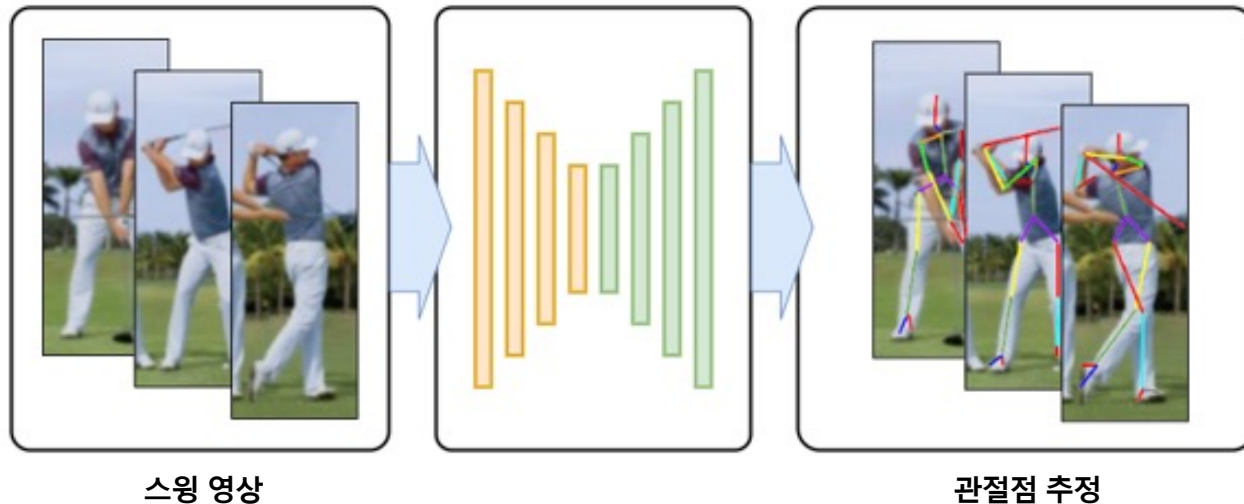
- 인체 관절점 및 클럽헤드를 추정하여 프레임 번호에 따른 좌표 메타데이터 리턴

■ 방법론

- 공개된 사람의 관절 정보를 포함한 데이터 셋을 활용하여 모델 학습 진행
- 제공된 골프 영상으로부터 클럽 헤드가 추가된 데이터로 모델 재학습 진행

■ 문제 정의

- 효율적인 데이터 어노테이션 정의 및 수행
- 클럽 헤드의 빠른 속도로 인한 추정 어려움
- 동영상 프레임마다 관절 좌표의 차이로 인한 떨림 개선

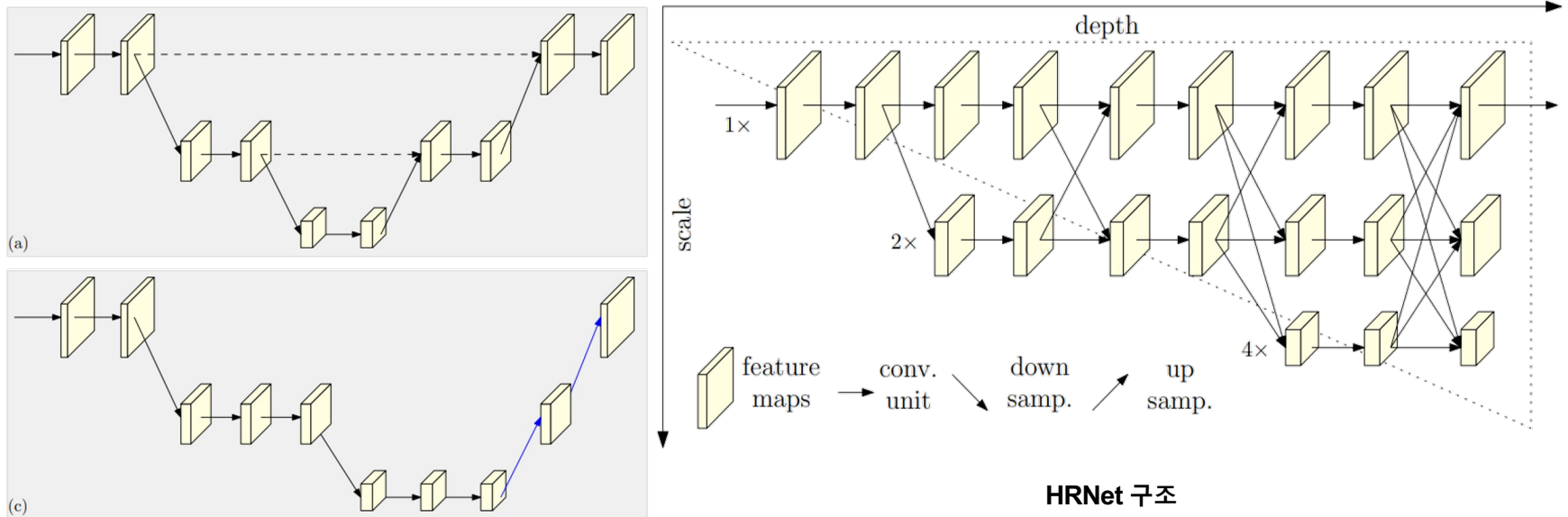


관절점 추정 프레임워크 (1/3)

■ 베이스라인 모델

➤ High-Resolution Network (HRNet)

- ✓ 고해상도의 정보를 유지하면서 저해상도의 피처를 병합하여 성능 향상을 이뤄냄
- ✓ 사람에게 해당 되는 영역만 잘라낸 후에 관절점을 추정하는 방식으로 골프 클럽 헤드를 포함하지 않기 때문에 바로 적용하기 어려움



관절점 추정 프레임워크 (2/3)

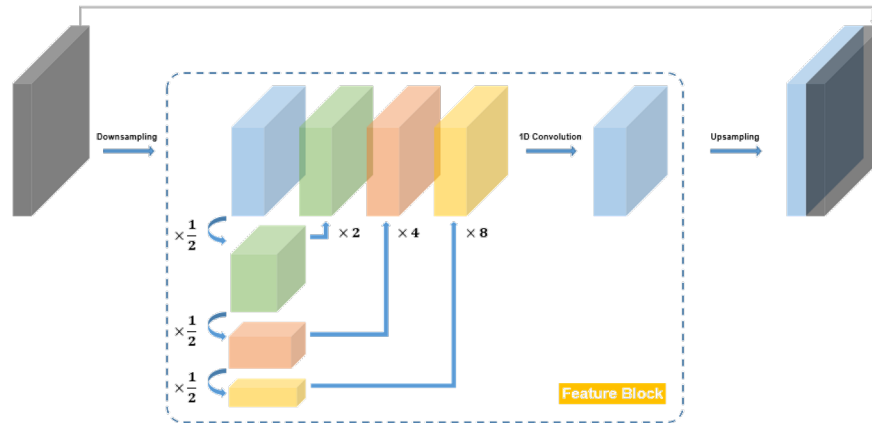
■ 제안 모델

➤ Scalable Pose Network

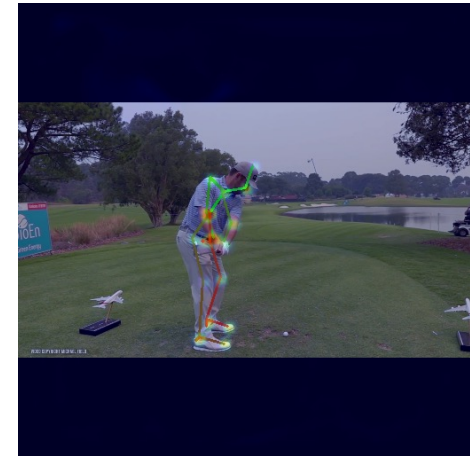
- ✓ 블록 단위의 피쳐 병합 구조를 만들어 여러 개의 스택을 쌓을 수 있는 네트워크 구조
- ✓ 원본 이미지의 해상도를 가지는 피쳐 사이즈를 유지하여 quantization error 줄임
- ✓ 학습 과정에서 각 블록 별로 heatmap을 추정하여 loss를 계산한 뒤, 앞에서 뒤로 갈수록 큰 loss를 가지도록 weighted-sum 하여 학습함
- ✓ 고해상도에서부터 저해상도까지 피쳐를 사용하는 범위나 네트워크의 사이즈를 자유롭게 조절 가능함
- ✓ 추정된 21개의 heatmap에서 가장 높은 확률값 만을 추출하여 최종 위치 추정



입력 이미지



Scalable Pose Network 구조

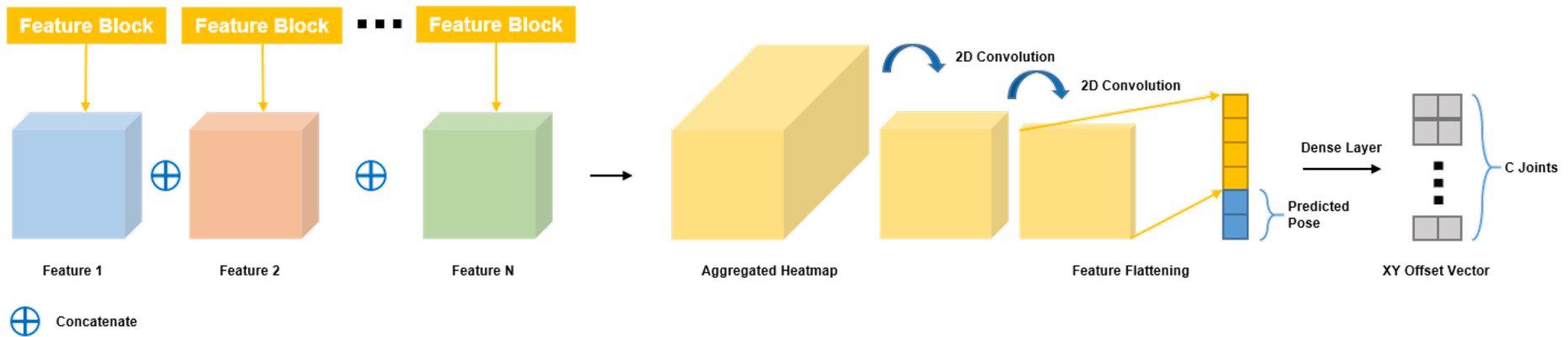


관절점 추정 결과

관절점 추정 프레임워크 (3/3)

➤ Refinement Network

- ✓ 관절점 추정 네트워크에서 얻어낸 관절점 위치를 세부적으로 조정하는 2차원 offset vector를 추정하는 네트워크 구조
- ✓ 각 블록별 heatmap을 병합하여, 2차원 Convolution을 통해 피쳐 벡터를 얻은 뒤 기존 관절점 위치와 함께 1번의 dense layer를 통과함
- ✓ 추정된 offset vector를 스케일링하여 기존에 예측한 관절점 위치에 더함
- ✓ 기존에 다운 샘플링된 heatmap에서 가장 높은 확률값을 가지는 픽셀 단위로 추정하면서 발생하는 quantization 에러를 offset vector regression을 통해 해결함

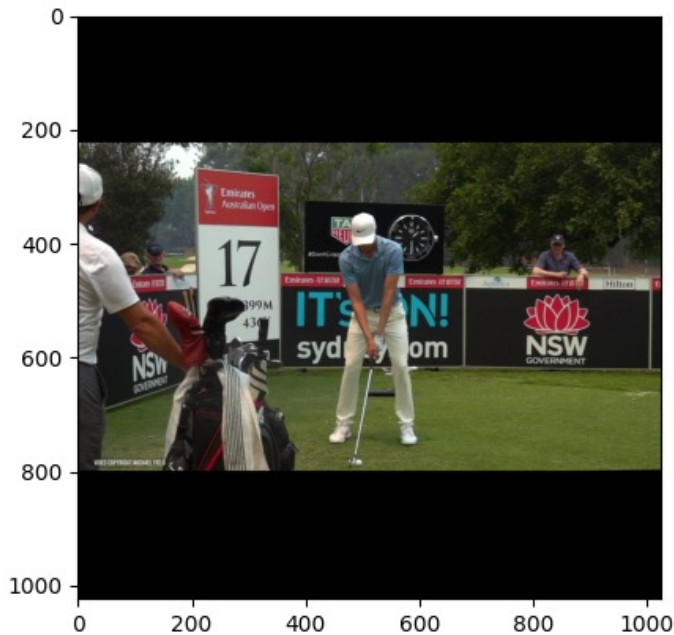


Refinement Network 구조

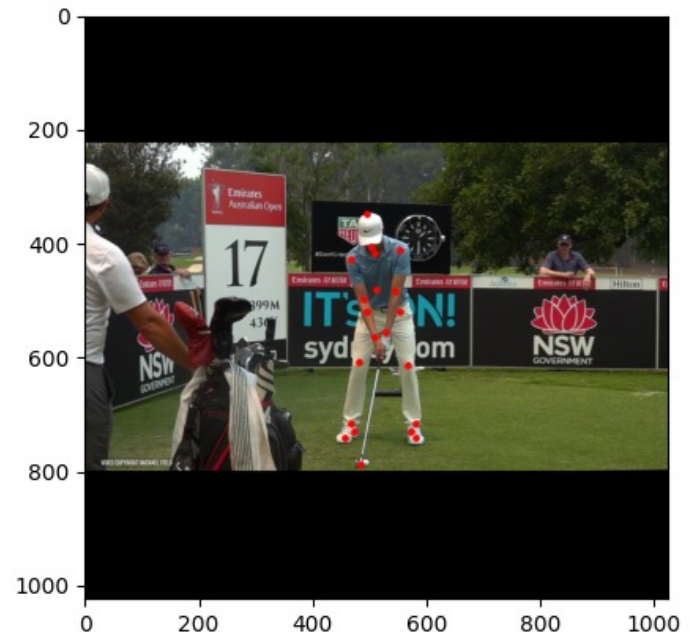
관절점 추정 프레임워크 학습 (1/2)

■ 학습 이미지와 레이블 예시

- 관절점에 해당하는 픽셀 위치 주변에 2D Gaussian heatmap 레이블을 생성
- 블러가 심한 골프 클럽 헤드를 잘 찾아내기 위해 가이드가 될 수 있는 골프채 영역을 따라 레이블을 추가하여 학습 진행
- 일정 확률에 따라 골프채 영역에 블러를 주어, 테스트 환경에서의 성능 향상



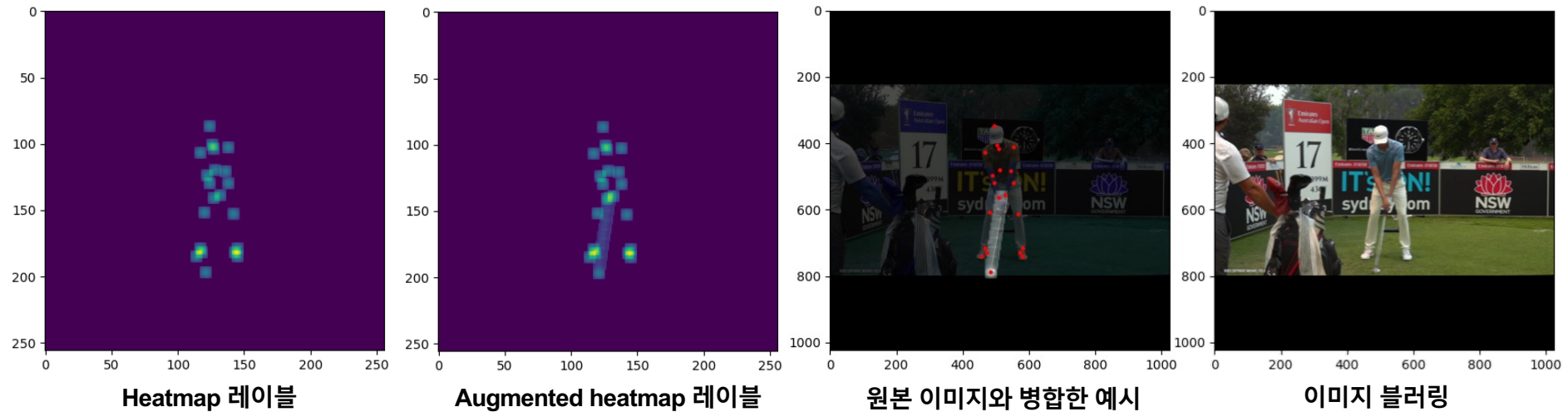
입력 이미지



레이블

관절점 추정 프레임워크 학습 (2/2)

■ 골프채 영역에서의 heatmap augmentation



■ 영상에서의 관절점 떨림 현상 원인 분석

- 추정된 heatmap 영역의 차이는 미세하지만, 가장 높은 확률값을 가지는 픽셀의 위치가 흔들림
- Upsampling 과정에서 quantization error 와 떨림 정도가 극대화됨

■ 프레임간의 정보 차이를 활용한 관절점 위치 세부 조정

Voice Caddie 데이터셋

- Train (80%): 32,000개 프레임
- Test (20%): 8,000개 프레임



	Label
0	Top Head
1	Nose
2	Neck
3	Chest
4	Right Shoulder
5	Right Elbow
6	Right Wrist
7	Left Shoulder
8	Left Elbow
9	Left Wrist
10	Right Hip
11	Right Knee
12	Right Ankle
13	Left Hip
14	Left Knee
15	Left Ankle
16	Left Big Toe
17	Left Heel
18	Right Big Toe
19	Right Heel
20	Golf Club Head

성능 평가

■ 평가 지표

➤ Percentage of Correct Keypoints head-normalized (PCKh@0.5)

➤ 목표 성능

✓ 85%

➤ 구현 성능

✓ 2_6_1024: 95.75% (초기 성능), 95.81% (정제 성능)

✓ 2_6_1024_blur: 95.95% (초기 성능), 95.59% (정제 성능)

Top Head	Nose	Neck	Chest	Right Shoulder	Right Elbow	Right Wrist
98.33	97.97	98.69	98.84	96.37	90.12	94.11
98.47	97.67	98.91	98.98	96.00	91.06	93.90
98.47	97.53	98.69	98.62	96.44	91.64	94.69
97.38	96.66	99.06	98.76	96.95	91.57	94.26

Left Shoulder	Left Elbow	Left Wrist	Right Hip	Right Knee	Right Ankle	Left Hip
95.86	91.06	92.88	98.47	98.84	98.69	96.29
95.93	91.28	93.31	98.69	99.06	98.76	96.80
95.49	91.72	93.75	98.11	98.62	98.91	96.51
95.71	91.79	93.68	98.40	98.84	98.76	96.80

Left Knee	Left Ankle	Left Big Toe	Left Heel	Right Big Toe	Right Heel	Golf Club Head
98.26	99.13	97.89	98.62	98.55	98.18	73.62
98.26	99.06	97.89	98.62	98.62	98.33	72.46
98.47	99.13	97.97	98.18	98.33	97.97	75.73
98.55	99.13	98.40	98.33	98.33	98.33	67.81

관절점 추정 결과 (3/5)



주제 2: 스윙 영상에서 주요 스윙시퀀스 프레임 추정 알고리즘 개발

■ 목표

- 8가지 스윙 시퀀스가 발견된 각각의 프레임 번호 리턴
 - ✓ 어드레스, 테이크백, 백스윙, 톱, 다운스윙, 임팩트, 팔로스루, 피니시

■ 방법론

- 주요 스윙 시퀀스의 프레임을 추정하는 딥러닝 기반 네트워크 구축
- 검출된 관절과 공의 위치를 활용하여 영상 처리 기반 기술 개발

■ 성능평가

- 8가지 스윙 시퀀스에 대해 Percentage of Correct Events (PCE) 75% 달성



어드레스 테이크백 백스윙 톱 다운스윙 임팩트 팔로스루 피니시

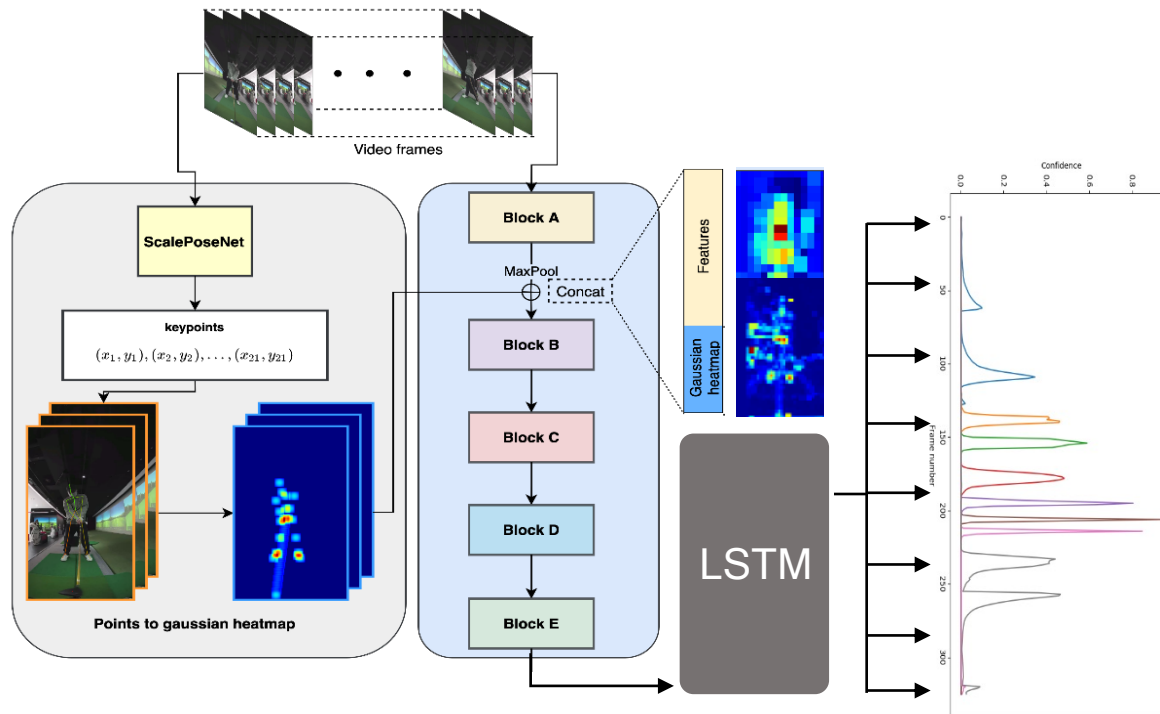
번호	스윙 시퀀스
1	어드레스
2	테이크백
3	백스윙
4	톱
5	다운스윙
6	임팩트
7	팔로스루
8	피니시

스윙시퀀스 프레임 추정 프레임워크

제안 모델

➤ Pose-Guided SwingNet (PGSwingNet)

- ✓ 각 프레임 피쳐에 액션의 근거가 될 수 있는 포즈 정보를 가이드 정보로 함께 제공
- ✓ 학습 과정에서 손실 되는 정보를 최소화하기 위해 Residual approach를 이용하여 각 프레임별 최종 피쳐를 생성함
- ✓ 프레임별 최종 피쳐들을 시간별로 LSTM에 통과시켜 공간정보를 시간정보별로 학습함
- ✓ 각 프레임별 각 액션의 확률을 추정하여 각 액션의 가장 높은 확률 값을 갖는 이미지 프레임을 각 스윙 시퀀스로 추정함



Action	frame_num
Address	108
Take-back	142
Back-swing	159
Top	178
Down-swing	197
Impact	201
Follow-throw	211
Finish	240

PGSwingNet 구조

성능평가

평가 지표

➤ Percentage of Correct Events (PCE)

➤ 목표성능

✓ 75%

➤ 구현 성능

✓ 91.09%

Result of refined method without pose-guided heatmap (%)

Address	Take-back	Backswing	Top	Downswing	Impact	Follow-through	Finish	Average PCE	Avg w/o AD & F
45.62	82.18	83.08	52.87	97.28	99.67	93.05	18.73	71.60	84.74

Result of proposed method (%)

Address	Take-back	Backswing	Top	Downswing	Impact	Follow-through	Finish	Average PCE	Avg w/o AD & F
50.45	83.69	89.12	80.97	96.98	99.70	96.07	40.48	79.68	91.09

주제 3: Annotation 작업 수행

■ 데이터 셋 구성

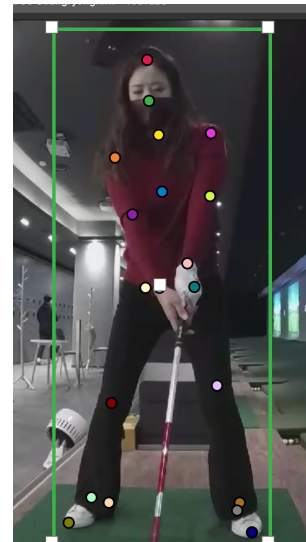
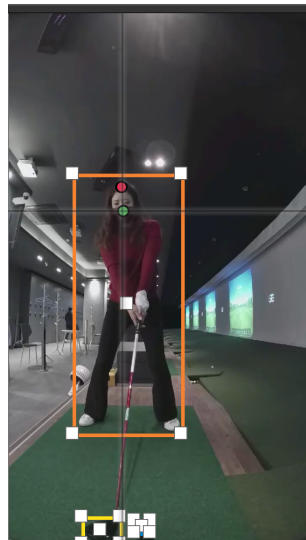
➢ 136,000 frames = 16,000 frames + 120,000 frames

✓ Annotation 수행 할 20 프레임 선별 : 16,000 frames = 2,000 sets x 8 frames

✓ B. box, seg, pose 수행 : 120,000 frames = 2,000 sets x 20 frames x 3 tasks

■ 동영상 내 20 프레임 선별 기준

➢ 동영상 전체를 검토한 후 작업해야 하므로 다른 작업에 비해 많은 시간이 걸림



■ 연구제목

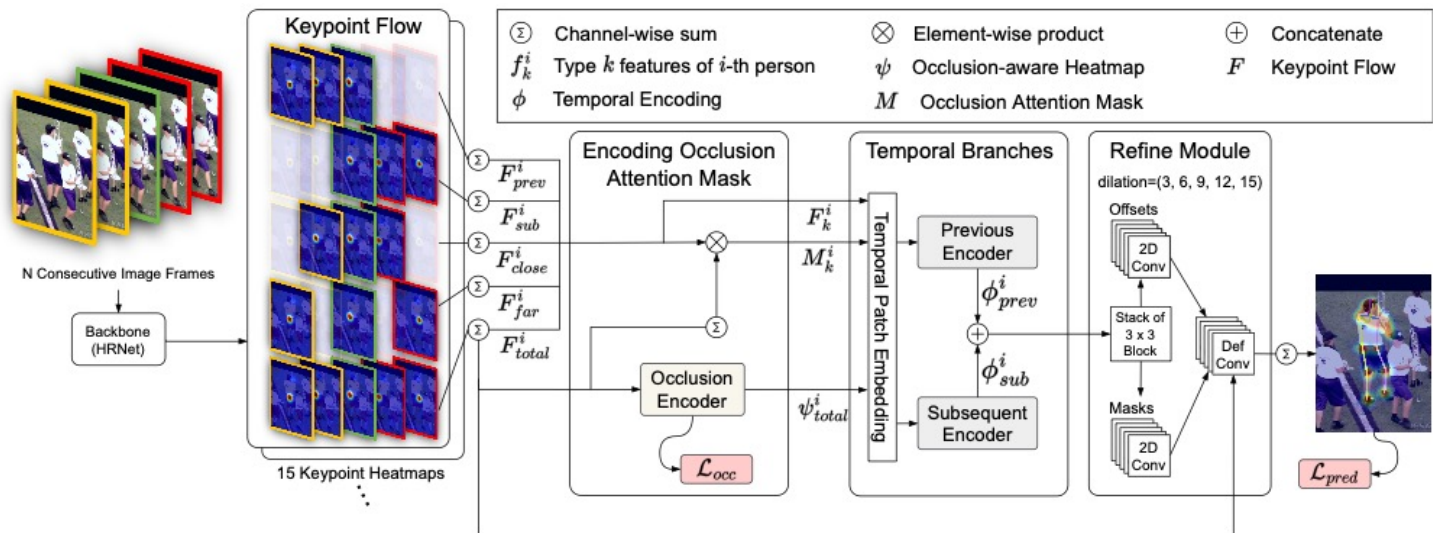
- OTPose: Occlusion-Aware Transformer for Pose Estimation in Sparsely-Labeled Videos

■ 연구목표

- 동영상 포즈추정 시 모든 프레임에 레이블을 달지 않은 데이터 이용
- Occlusion에 영향을 받지 않는 모델 개발

■ 방법론

- 관절 간 관계를 attention하는 시간 분기 Transformer branch
- Occlusion된 부위를 attention하는 mask를 제안해 명시적으로 학습

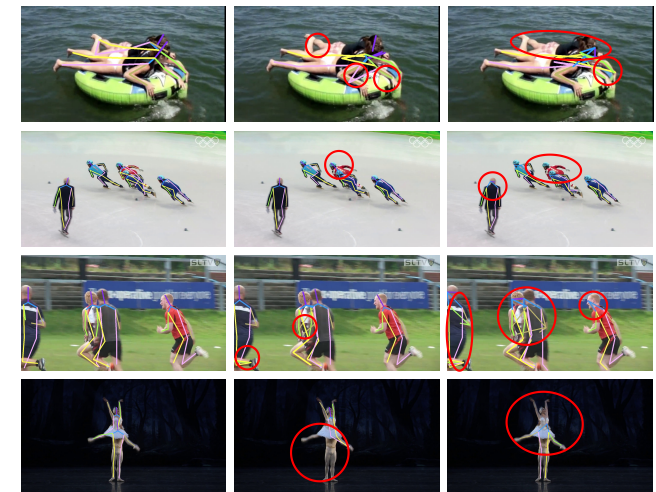


■ 평가지표

➤ mean Average Precision (mAP)

Result on PoseTrack17 dataset

Method	Head	Should.	Elbow	Wrist	Hip	Knee	Ankle	Mean
PoseFlow	66.7	73.3	68.3	61.1	67.5	67.0	61.3	66.5
SimpleBaseline	81.7	83.4	80.0	72.4	75.3	74.8	67.1	76.7
HRNet	82.1	83.6	80.4	73.3	75.5	75.3	68.5	77.3
MDPN	85.2	88.8	83.9	77.5	79.0	77.0	71.4	80.7
PoseWarper	81.4	88.3	83.9	78.0	82.4	80.5	73.6	81.2
DCPose	88.0	88.7	84.1	78.4	83.0	81.4	74.2	82.8
OTPose	90.7	91.5	86.5	80.2	85.2	84.6	80.5	86.0



OTPose

DCPose

HRNet

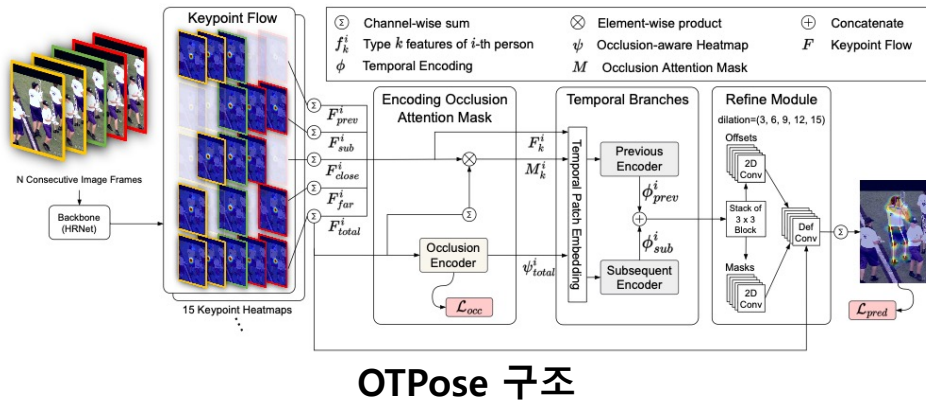
미소 인공지능 모델 개발 챌린지 대상

주제

- 영유아 행동 영상 데이터 부문
- 연구기반 모델로 대상 수상

주관

- 과학기술정보통신부·한국지능정보사회진흥원 지원, (주)미소정보기술



주최 과학기술정보통신부 NIA 한국지능정보사회진흥원 주관 (주)미소정보기술

챌린지 주제

10월 7일부터 11월 11일까지 참가 접수를 받아 진행된 '2021 미소 인공지능 모델 개발 챌린지'는 인공지능 개발자, 빅데이터 관련 기업 및 연구기관 종사자, 대학·대학원생 등 만 15세 이상을 대상으로 행사를 진행하였다. ▲ 팀 역량 ▲ AI 개발 경험 ▲ 인공지능 모델 개발 계획 등을 기준으로 한 예선 심사를 통해 10개의 팀이 선정되었으며, 2주간 아이디어 및 AI 모델 개발 공모를 진행하였다.

본 행사의 대상은 KKJY 팀(고려대학교 인공지능학과 소속 양건준, 김현우, 진경민, 강태경), 최우수상에는 AI고난 팀(서울대학교 전기정보공학부 정보신호처리 연구실 소속 이성주, 박선지, 김민우, 성균관대학교 미디어 시스템연구실 소속 이성구), 우수상에는 BrianaAI 팀(인피티니 헬스케어 소속 변소현)이 선발되었다. 시상식은 온라인으로 진행되었으며, 상장과 대상 300만 원, 최우수상 150만 원, 우수상 50만 원의 상금이 주어졌다.

기사 내용 발췌

■ 연구제목

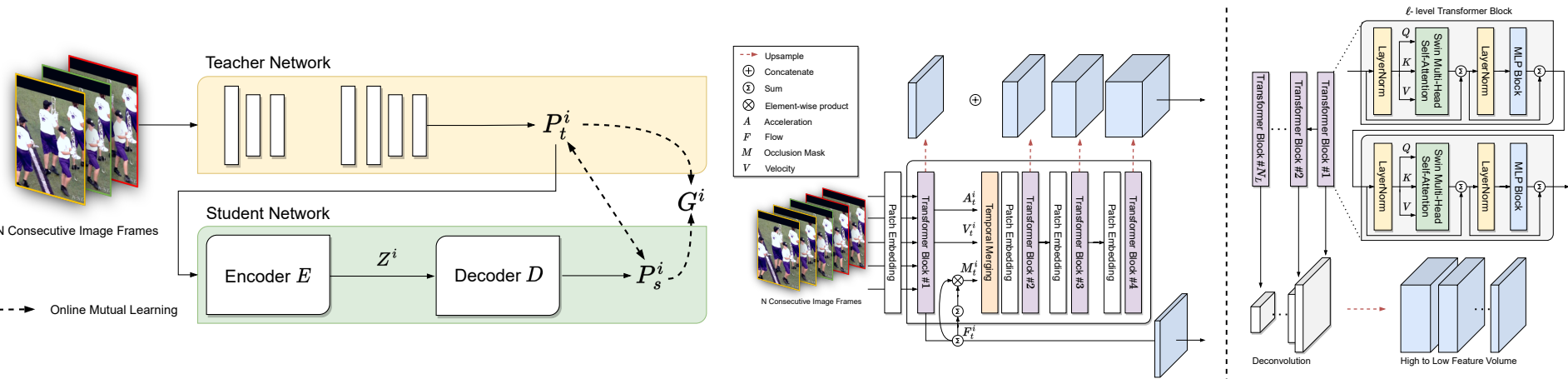
- Exploiting Keypoint Motion for Human Pose Estimation with Scaled Transformer

■ 연구목표

- 동영상 포즈 추정에서 백본 네트워크의 영향을 받지 않고 Occlusion, jitter 해결

■ 방법론

- 관절 움직임 정보를 계산해 Occlusion, jitter 문제완화
- 여러 해상도의 heatmap을 사용해 정확도 개선
- 백본 네트워크와 제안 네트워크 상호 학습 (Online Mutual Learning)



감사합니다.