


SwingNetV2



1조 서상원 안지현 임연오
2025.06.18



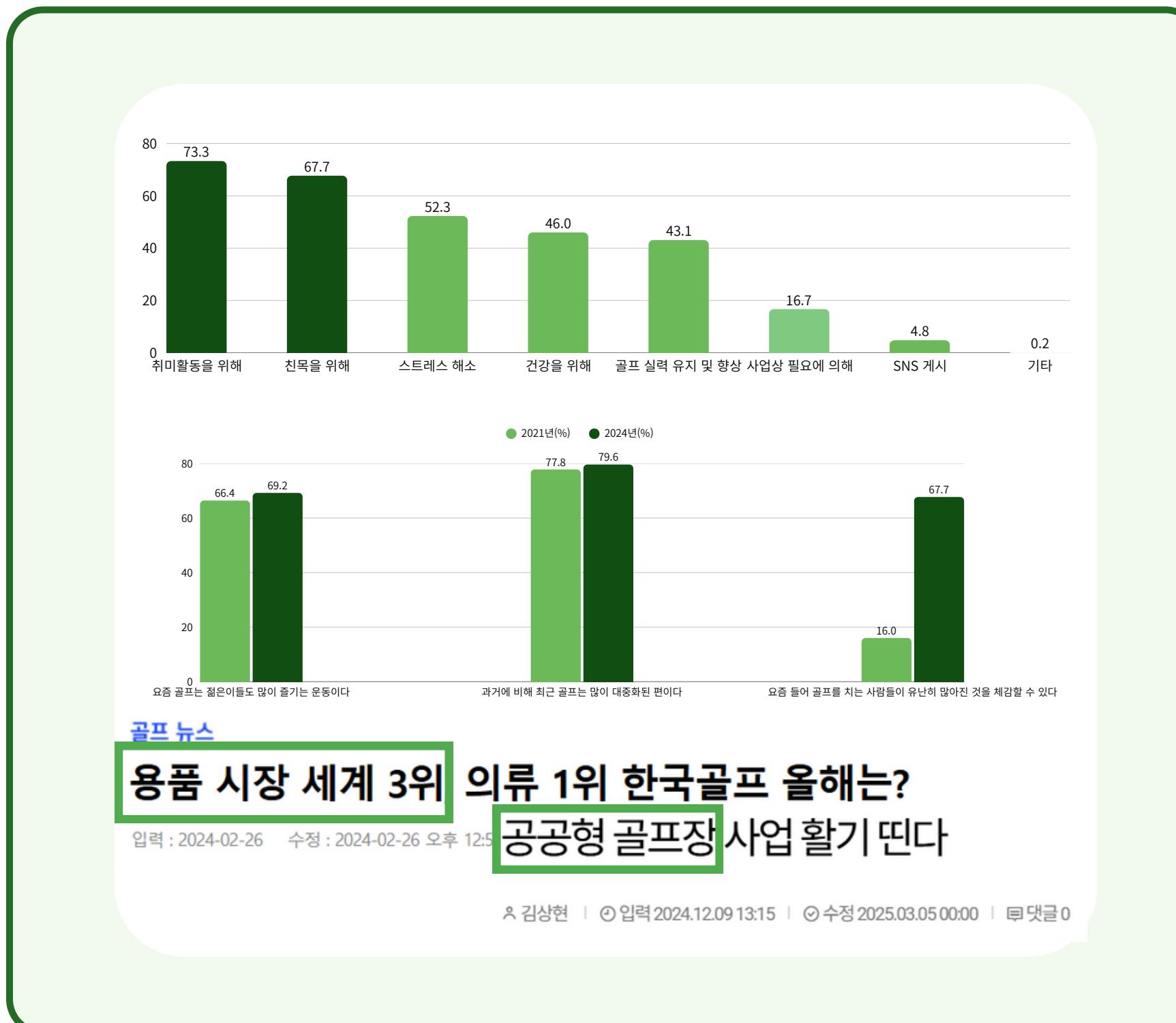
목차

● 서비스 소개 -----	3
● 데이터 -----	15
● 모델 개발 -----	21
● 검증 및 테스트 -----	35
● 기대효과 -----	38



서비스 소개

골프의 대중화



- **국내 골프 용품 시장 미국, 일본에 이어 세계 3위 수준**
- **정부, 2027년까지 공공형 골프장 30개소 확충 계획**
- **골프 활동 목적 1,2위는 취미 활동 및 친목을 위한 것으로 골프의 생활체육화 활발**
- **성인 10명 중 8명(79.6%)이 “골프가 과거보다 많이 대중화됐다”고 응답**

골프 스윙 자세의 중요성



Low back pain and golf: A review of biomechanical risk factors

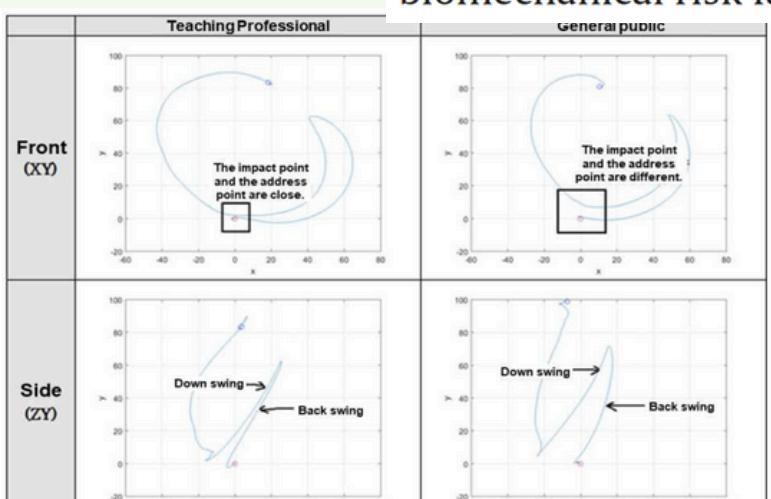


그림 6. 티칭프로(좌)와 일반인(우)의 스윙 플레인
Fig. 6. Swing planes of teaching pros(L) and public(R)

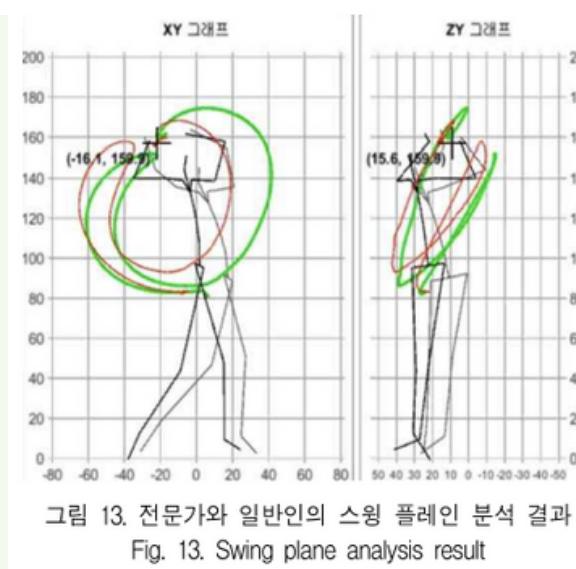


그림 13. 전문가와 일반인의 스윙 플레인 분석 결과
Fig. 13. Swing plane analysis result

- 골프 손상의 가장 흔한 원인에는 과도한 훈련 및 연습, 나쁜 스윙 메카닉스, 관절과 근육의 유연성 결핍, 불충분한 워밍업 등

대한스포츠의학회

- 비거리에 중대한 영향을 미치는 스윙플레인 궤적에서 **프로와 일반인의 차이가 두드러짐**

고경리, 반성범, "관성센서 기반의 골프 스윙 모션 분석," Journal of KIIT, 16(1), 125-131, 2018

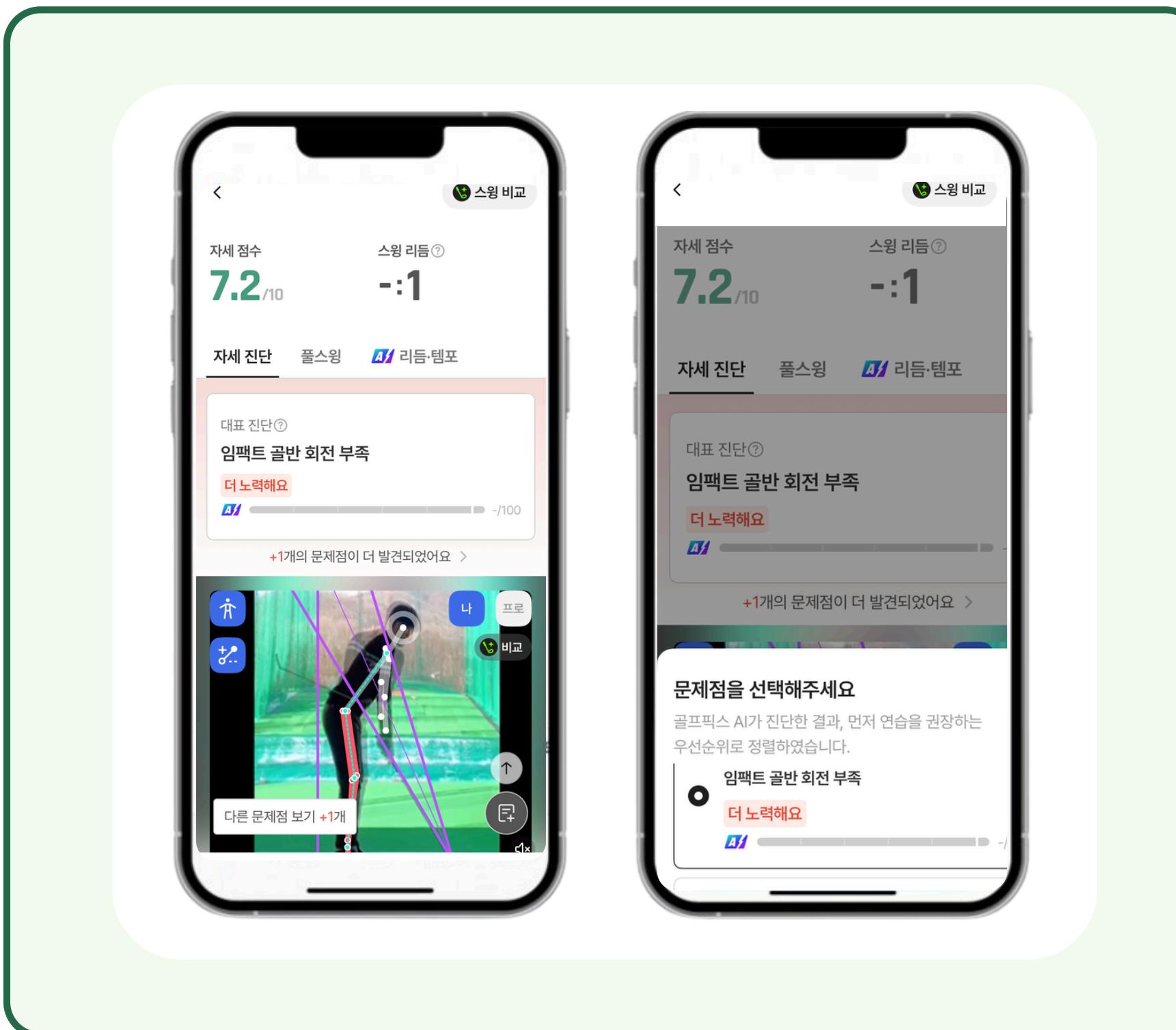
- **this risk is further inflated when a golfer exhibits poor technique. Forces capable of causing injury are generated during the golf swing, and if not properly sequenced, these forces dissipate into joints and soft tissue increasing the risk of injury.**

Nathan Edwards, Clark Dickin, Henry Wang, "Low back pain and golf: A review of biomechanical risk factors," Journal of Sport and Health Science, 2020

사용자의 스윙 단계별 자세를 분석하여
피드백을 제공하는 유사 서비스



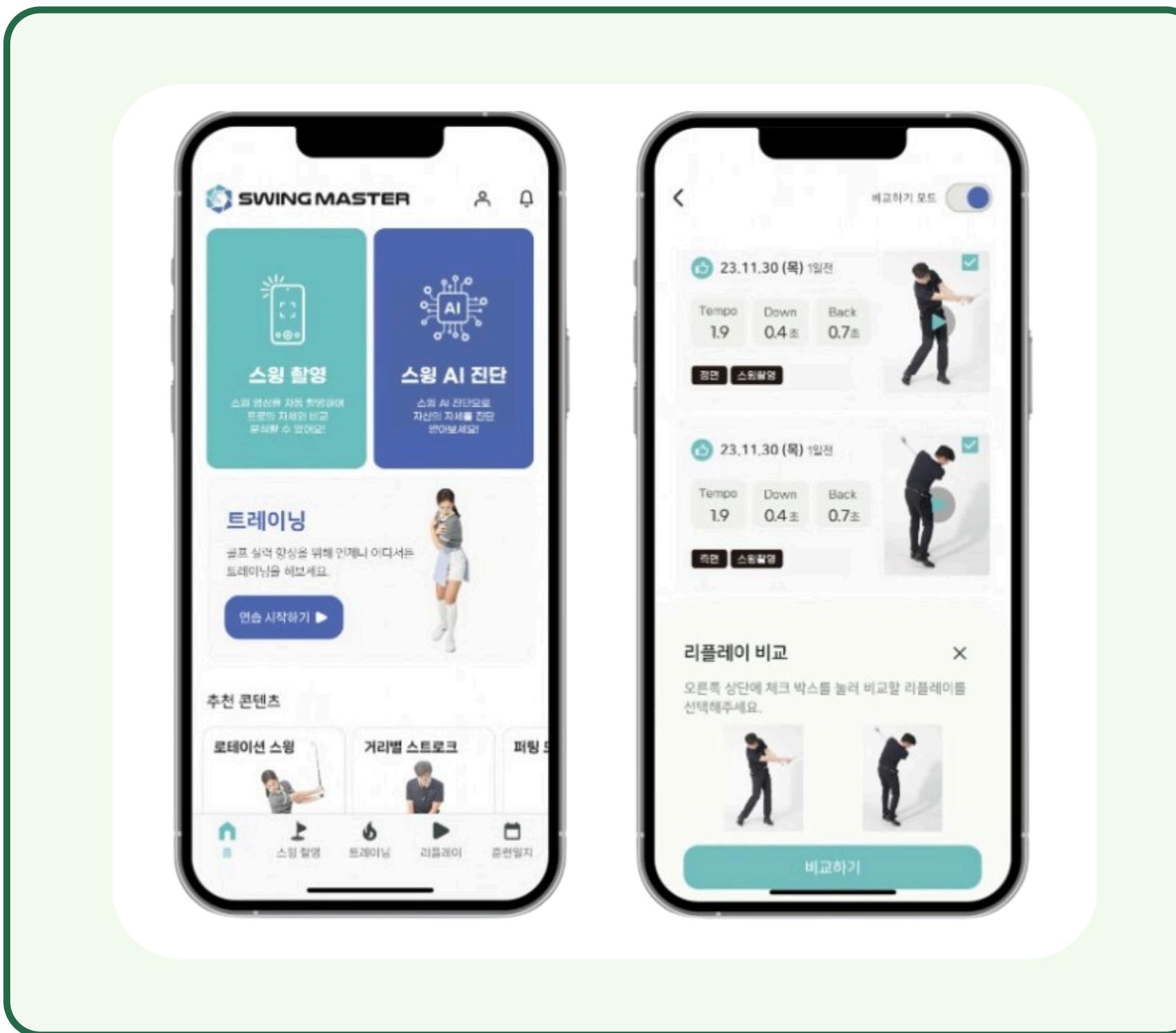
골프픽스



- 국내 누적 200만+ 다운로드, 사용자 수 기준 1위
- 2D AI 포즈 추정 기반 스윙 분석
- 프로 스윙 비교 수준별 진단
- 맞춤형 드릴 영상 추천 및 학습 플랜 제공

AI 분석 결과와 실제 스윙 평가 점수의
신뢰도가 낮아, 사용자마다 점수 편차가
크고 객관적 기준이 부족함

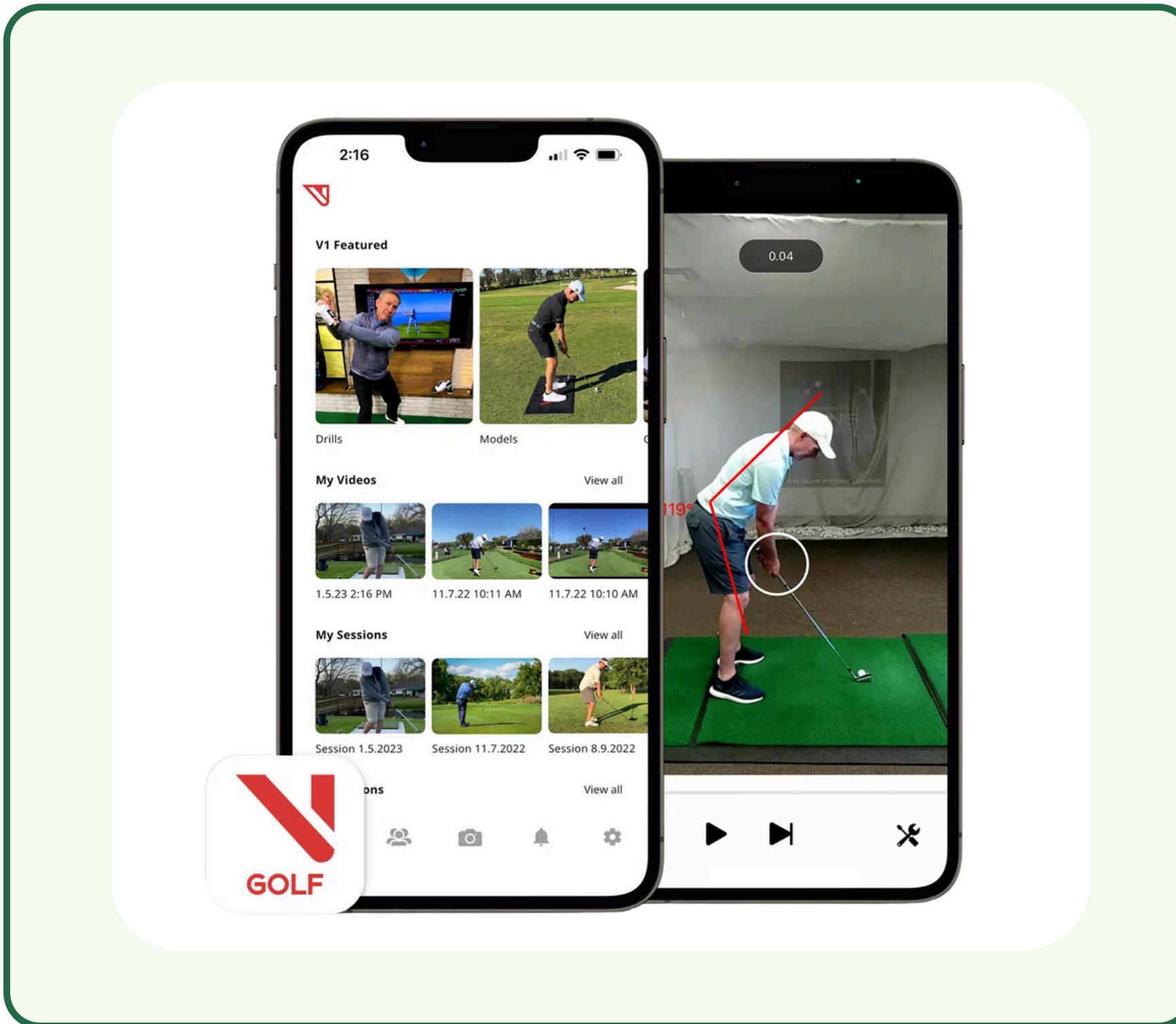
스윙마스터



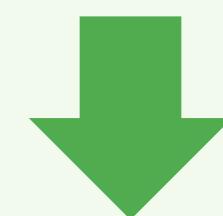
- 실시간 즉각 피드백
- 프로의 전문화된 코칭 팁 제공
- 스윙 영상·음성 입력을 포함한 훈련일지
- 실제 골프 프로가 콘텐츠 제작에 참여

AI 분석과 피드백이 제공되지만, 자세 평가와 점수 산정이 여전히 주관적 요소에 의존하여 일관된 객관성 확보에 한계가 있음

V1 GOLF



- 드로잉 도구(라인·각도)로 관절·스윙 포인트 정확 측정
- 프로 모델 스윙과 오버레이 비교로 자세 차이 분석
- 영상 뒤집기 기능으로 좌·우타 모두 직관적 확인

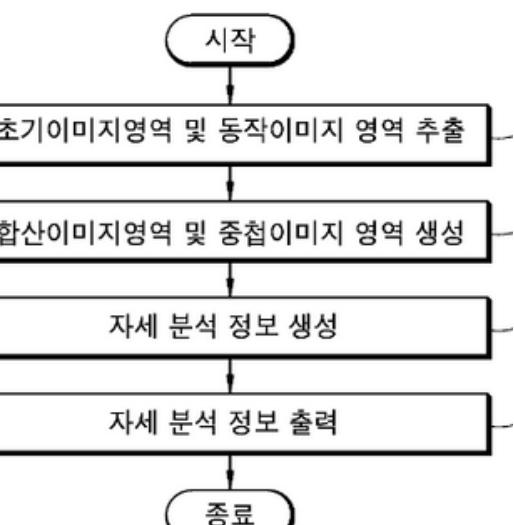
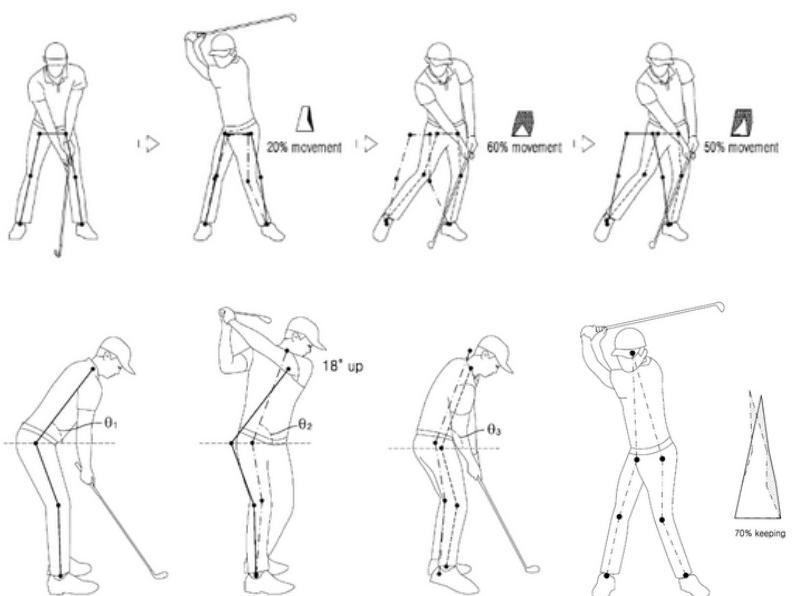


사용자의 해석이나 코치의 주관에 크게
의존해 객관적 지표로 활용하기 어려움

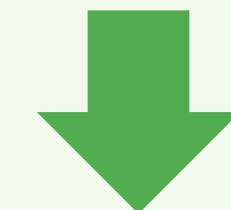
특허

전체 청구항 수 : 총 20 항

(54) 발명의 명칭 골프 스윙에 따른 자세 분석장치 및 방법



- 영상에서 추출한 초기 자세와 스윙 자세 이미지를 합산
- 두 이미지의 공통부와 합산부를 시각화하여
인체 자세 변화를 정량적으로 분석하는 방법을 제안



머신러닝·딥러닝 X
이미지 처리 기법(ROI 추출, 픽셀 연산)
기반 시스템

AI Golf: Golf Swing Analysis Tool for Self-Training

IEEE Access
Multidisciplinary | Peer-reviewed | Open-access journal

Received 20 August 2022, accepted 18 September 2022, date of publication 6 October 2022, date of current version 11 October 2022.
Digital Object Identifier 10.1109/ACCESS.2022.3210260

RESEARCH ARTICLE

AI Golf: Golf Swing Analysis Tool for Self-Training

CHEN-CHIEH LIAO^{✉1}, DONG-HYUN HWANG², AND HIDEKI KOIKE¹
¹Department of Computer Science, Tokyo Institute of Technology, Tokyo 152-8550, Japan
²NAVER CLOVA Voice&Avatar, Seongnam 13581, South Korea
Corresponding author: Chen-Chieh Liao (liao.c.cs@mit.ac.jp)
This work was supported by the Japan Science and Technology Agency (JST) Core Research for Evolutional Science and Technology (CREST), Japan, under Grant JPMJCR17A3.

ABSTRACT In the field of the acquisition of sports skills, a common way to improve sports skills, such as golf swings, is to imitate professional players' motions. However, it is difficult for beginners to specify the keyframes on which they should focus and which part of the body they should correct because of inconsistent timing and lack of knowledge. In this study, a golf swing analysis tool using neural networks is proposed to address this gap. The proposed system compares two motion sequences and specifies keyframes in which significant differences can be observed between the two motions. In addition, the system helps users intuitively understand the differences between themselves and professional players by using interpretable clues. The main challenge of this study is to target the fine-grained differences between users and professionals that can be used for self-training. Moreover, the significance of the proposed approach is the use of an unsupervised learning method without prior knowledge and labeled data, which will benefit future applications and research in other sports and skill training processes. In our approach, neural networks are first used to create a motion synchronizer to align motions with different phases and timing. Next, a motion discrepancy detector is implemented to find fine-grained differences between motions in latent spaces that are learned by the networks. Furthermore, we consider that learning intermediate motions may be feasible for beginners because, in this way, they can gradually change their pose to match the ideal form. Therefore, based on the synchronization and discrepancy detection results, we utilize a decoder to restore the intermediate human poses between two motions from the latent space. Finally, we suggest possible applications for analyzing and visualizing the discrepancy between the two input motions and interacting with the users. With the proposed application, users can easily understand the differences between their motions and those of various experts during self-training and learn how to improve their motions.

INDEX TERMS Computer vision, machine learning, motor skill training, golf.

I. INTRODUCTION
In sports, it is difficult for beginners to improve their skills without prior knowledge or assistance from coaches. As a conventional method, people go to lessons to meet experts and learn how to play in the proper form. However, in most sports, players spend considerable time training alone to achieve outstanding results in the field and retain exceptional body conditions. Therefore, it is important to design and implement an effective and accurate self-training process for such situations.

In the field of the acquisition of sports skills, one way to improve sports skills is to replicate professional players' motions. People watch the movements of professional athletes on television or the Internet and try to make their bodies move similarly to professionals. To accelerate this process, many systems have recently been developed to help users understand the movements of professionals [1], [2], [3], [4], [5]. However, in these previous works and systems, users may struggle to refine their movement with no idea which timing of the whole motion, which parts of the body they should focus on, or how they can change their body movements to get their form closer to that of professionals.

With significant advances in machine learning technologies, many systems have been built to recognize different objects, make predictions for decisions, or even predict the future [6]. Researchers have focused on producing self-training systems with neural networks [7]. A recent study [8] introduced a climbing training system in which users can receive recommended poses and movements to

The associate editor coordinating the review of this manuscript and approving it for publication was Jinja Zhou.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see <https://creativecommons.org/licenses/by/4.0/>. VOLUME 10, 2022

106286

IEEE 2022, 피인용수 약 35회

연구 목표

- **셀프 어텐션 기반 TCC(Temporal Cycle-Consistency)**
네트워크를 활용하여, 사용자가 자신의 골프 스윙과 전문가의 스윙을 비교하고, 중간 자세 생성을 통해 이상적인 폼을 학습할 수 있는 AI 기반 자체 훈련 도구 개발

모델 구조: TCC 네트워크 기반으로 세 가지 모듈 구성

- **V-TCC:** 비디오 입력 처리
- **S-TCC:** 3D 인간 자세 데이터 활용
- **SA-TCC:** 신체 부위별 중요도 학습

성능 및 기대

- **성능:** SA-TCC 모델이 불일치 감지에서 최고 성능 달성 (상관계수 0.76 이상)
- **기대:** 사용자가 코치의 도움 없이 시각적 피드백과 중간 자세 안내를 통해 스윙을 효과적으로 교정 가능

GolfPose: Golf Swing Analyses with a Monocular Camera Based Human Pose Estimation

GOLFOSE: GOLF SWING ANALYSES WITH A MONOCULAR CAMERA BASED HUMAN POSE ESTIMATION

Zhongyu Jiang^{1*}, Haorui Ji^{2*}, Samuel Menaker² and Jeng-Neng Hwang¹

¹Dept. Electrical & Computer Engineering , University of Washington
²SPORTSBOX.AI INC.
zyjiang@uw.edu, haoruiji@sportsbox.ai, samm@sportsbox.ai, hwang@uw.edu

ABSTRACT

With the rapid developments of computer vision and deep learning technologies, artificial intelligence takes a more and more important role in sports analyses. In this paper, to attain the objective of automated golf swing analyses, we propose a lightweight temporal-based 2D human pose estimation (HPE) method, called GolfPose, which achieves improved performance than the state-of-the-art image-based HPE methods. Unlike traditional image-based methods, our temporal-based method, designed for efficient and effective golf swing analyses, takes advantage of the temporal information to improve the estimation accuracy of fast-moving and partially self-occluded keypoints. Furthermore, in order to make sure the golf swing analyses can run on mobile devices, we optimize the model architecture to achieve real-time inference. With around 10% of the parameters and half of the GFLOPs used in the state-of-the-art HRNet, our proposed GolfPose model can achieve 9.16 mean pixel error (MPE) in our golf swing dataset, compared with 9.20 MPE for HRNet. Furthermore, the proposed temporal-based method, facilitated with golf club detection(GCD), significantly improves the accuracy of keypoints on the golf club from 13.98 to 9.21 MPE.

Index Terms— Sports Analysis, Human Pose Estimation, Golf Swing, Line Segment Detection

1. INTRODUCTION

Sports are social-cultural activities and have already become important parts of our daily life. It allows people to interact with each other regardless of their social status, helps to improve the quality of people's lives, and also serves as a significant symbol for measuring the development and progress of a country and society.

A significant amount of resources have been allocated to the modern sports industry, which demands higher requirements not only on the athletes themselves but also in a lot of related supporting technologies. For example, tracking players' trajectories in the field[1] can improve the audiences' experience during game broadcasting, analyze and assess players' performance for better coaching[2], detect and prevent life-threatening situations to players, etc. These requirements call for accurate analyses of actions, conditions, and environments across different players, scenarios, and sports events.

Before the modern sports industry era, sports analytics could only use naked human eyes and their own experience to measure and analyze, which are unreliable, inefficient, and too subjective to generalize to different players, scenarios, and sports events. Therefore, manual analyses of sports are being gradually replaced by a combination of different sensors and algorithms that can automatically do all the cumbersome analyses. These capacities can help better assess the crucial sports event moments, resulting in more precise, efficient, and generalizable analysis results.

Among these newly emerging technologies, rapid development in computer vision communities combined with recent deep learning technologies have been highly appreciated in terms of efficiency and accuracy. Furthermore, thanks to the popularity of social media and online streaming, massive video and image data are generated and become available for researchers to utilize and improve the performance of their applications.

* These two authors contribute equally

More than 24.8 million people played golf in the U.S. in



Fig. 1. Two sample images from our dataset. The right one is annotated with 38 keypoints.

IEEE 2022, 피인용수 약 20회

연구 목표

- 모노클러 카메라만으로 3D 골프 스윙 궤적 분석
- 실시간 처리 가능한 경량화 아키텍처 제공
⇒ 모바일 환경 최적화

기술적 차별성

- **Golf Club Detection(GCD): 모션 블러/프레임 이탈 문제 해결을 위해 전통적 CV 기법과 딥러닝 결합**
- **2D → 3D 변환: 단일 뷰 영상에서 3D 스윙 궤적 복원 신체 관절(14개) + 클럽 키포인트(3개) 동시 추정**
- **시계열 최적화: 연속 프레임 간 temporal coherence 보장을 위한 LSTM 기반 포즈 보정**

성능 및 한계

- **정확도: 플레이어가 고정 위치에서 스윙하는 시나리오에 한정**
- **한계: 클럽 hosel 인식 실패 시 전체 분석 신뢰도 하락**

GolfPose: From Regular Posture to Golf Swing



연구 목표

- 일반 RGB 카메라로 골프 스윙 시 골퍼와 클럽의 2D/3D 자세를 동시 추정하는 GolfPose 모델을 제안

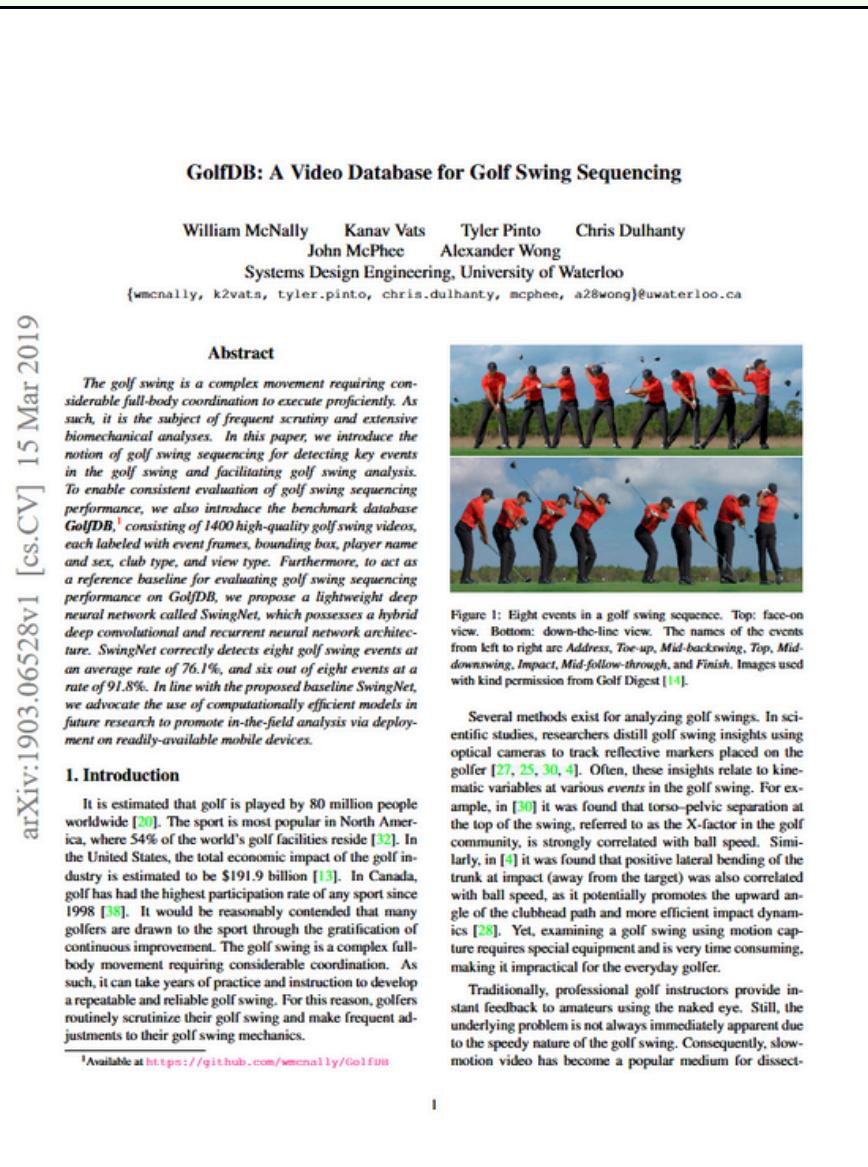
핵심 기술

- **GolfSwing 데이터셋:** Vicon 모션 캡처로 구축한 3D 골퍼(17-keypoint)-클럽(5-keypoint) Ground Truth 와 RGB 영상 동기화
- **모델 개선:** HRNet, ViTPose 등 2D 모델과 MixSTE 3D 리프팅 모델을 미세 조정해 정확도 향상
- **객체 탐지:** YOLOX-s 기반 통합 탐지로 실시간 성능 달성

성능 및 한계

- 클럽 정보 포함 시 3D 오차 32.3mm로 기존 대비 70% 개선, 부분 가림 상황에서도 강건성 확인
- 클럽 빠른 움직임으로 인한 흐림 효과가 주요 한계로 남음

GolfDB: A Video Database for Golf Swing Sequencing



CVPR 2019, 피인용수 약 80회

연구 목표

- 골프 스윙 비디오에서 8가지 핵심 이벤트를 프레임 수준으로 자동 검출하는 경량 모델인 **SwingNet** 제안

모델 구조

- SwingNet은 MobileNetV2 기반 CNN으로 공간 특징 추출
- Bi-LSTM으로 시계열 의존성을 학습하는 경량 하이브리드 딥러닝 모델
⇒ 모바일 디바이스에서도 실시간 분석 가능

성능

- 8개 이벤트 평균 정확도 **76.1 %**
- 핵심 6개 이벤트(Address · Finish 제외) 정확도 **91.8 %**

실무 활용성

- 실시간 피드백 앱, 원포인트 레슨
- 선수 퍼포먼스 모니터링 등 즉각적 활용 가능



데이터

GolfDB

GolfDB

Introduced by McNally et al. in [GolfDB: A Video Database for Golf Swing Sequencing](#)

GolfDB is a high-quality video dataset created for general recognition applications in the sport of golf, and specifically for the task of golf swing sequencing.

Source: [GolfDB: A Video Database for Golf Swing Sequencing](#)



[그림1] Face-on View



[그림2] Down-the-line View

- 데이터 규모: 남녀 프로 골퍼 1,400명의 HD 골프 스윙 영상으로 구성되며, 실시간 촬영 영상과 슬로우 모션 영상이 혼합
- 총 39만 여 프레임, 8개 이벤트 라벨 제공
- 골프 스윙 시퀀스에서 8가지 주요 동작 이벤트:
Address, Toe-up, Mid-backswing, Top, Mid-downswing, Impact, Mid-follow-through, Finish
- View 타입: Face-on & Down-the-line
⇒ 다양한 각도·클럽·성별 포함
- 추가 메타 정보: fps, 클럽 종류, 성별 등
⇒ 세분화 분석·도메인 적합성 ↑

GolfDB

8가지 이벤트 라벨



Address

Toe-up

Mid-
backswing

Top

Mid-
downswing

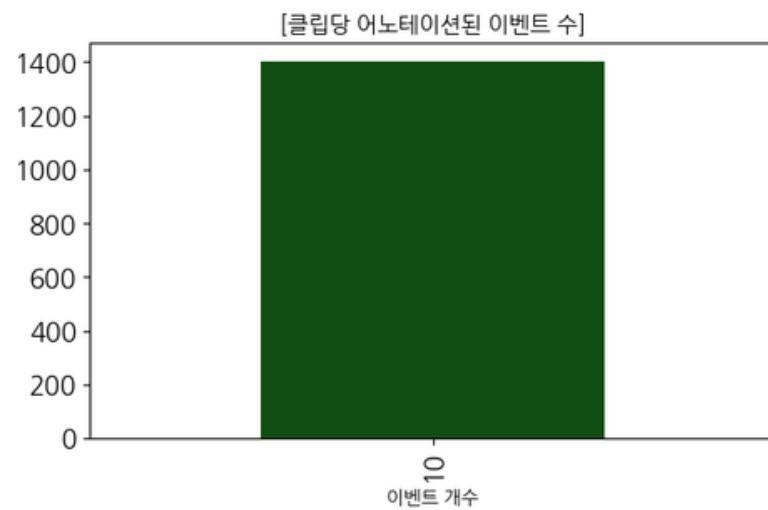
Impact

Mid-follow-
through

Finish

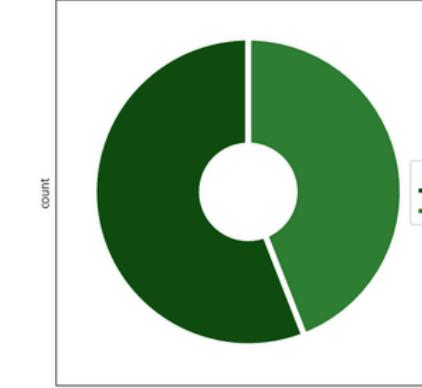
GolfDB EDA_1

1

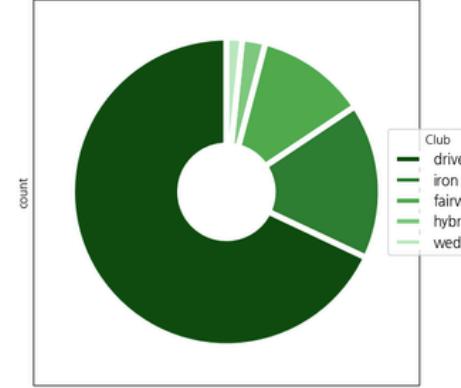


2

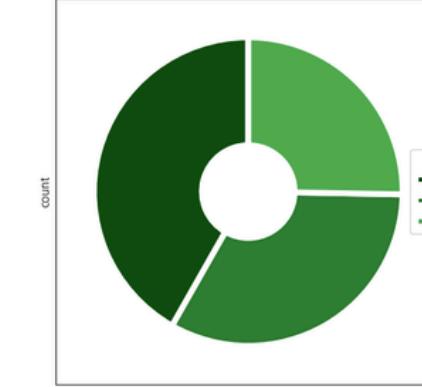
[선수 성별 (남/여)]



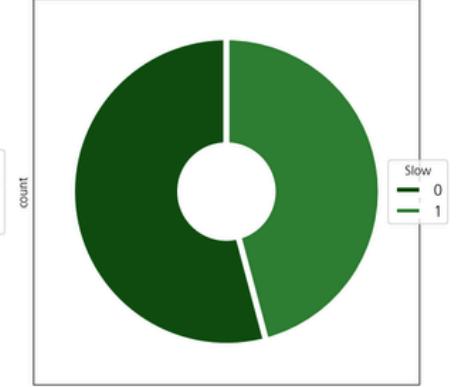
[사용된 Club 유형]



[촬영 view 유형]



[비디오 속도 (실시간/슬로우)]



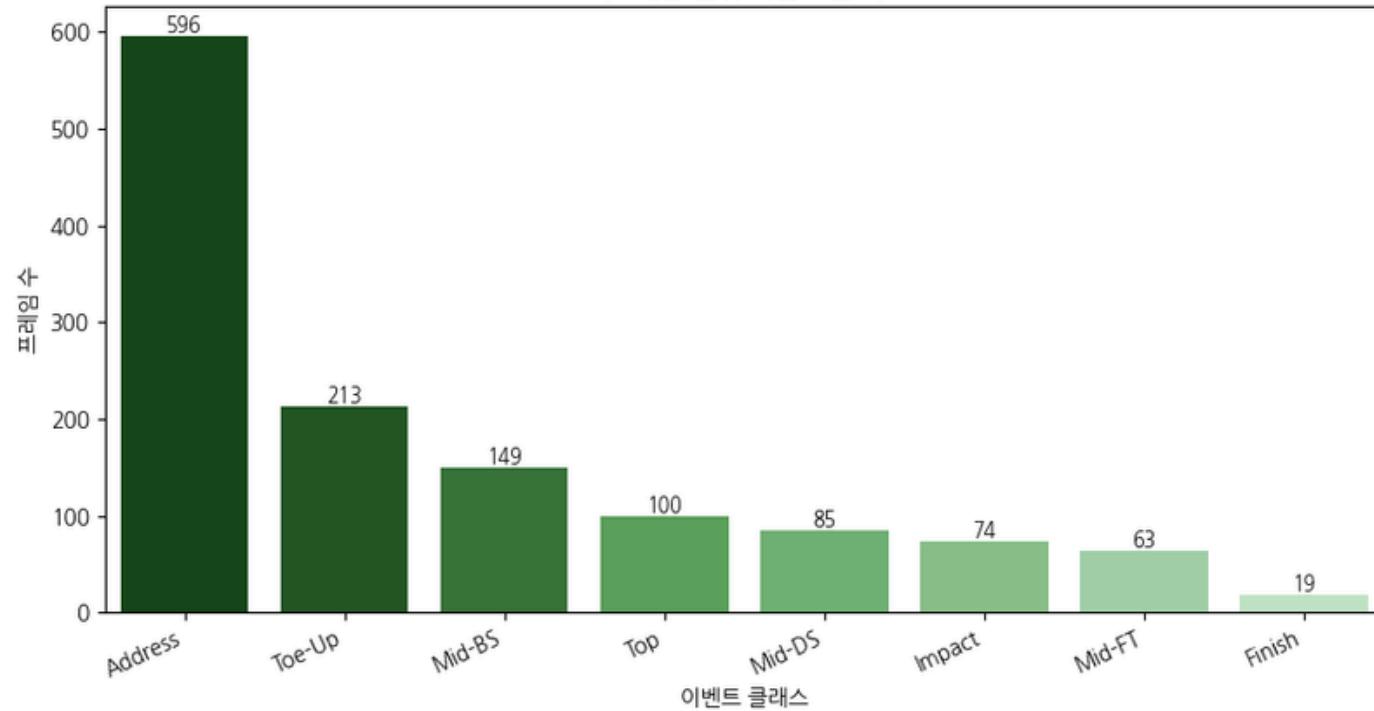
GolfDB: 총 580개 유튜브 영상에서 1,400개 스윙 클립을 추출함
각 클립마다 10개의 이벤트가 어노테이션되어 있음
→ 8개의 골프 스윙 이벤트 + 샘플의 시작, 끝 부분

- 성별: 남 55 % | 여 45 % → 분포 균일
- Club: 드라이버가 68 %, 아이언 16 %
- View: 다양한 각도 확보, 단 Down-the-line 우세
- 속도: 실시간 54 % | 슬로우 46 % → 분포 균일

GolfDB EDA_2

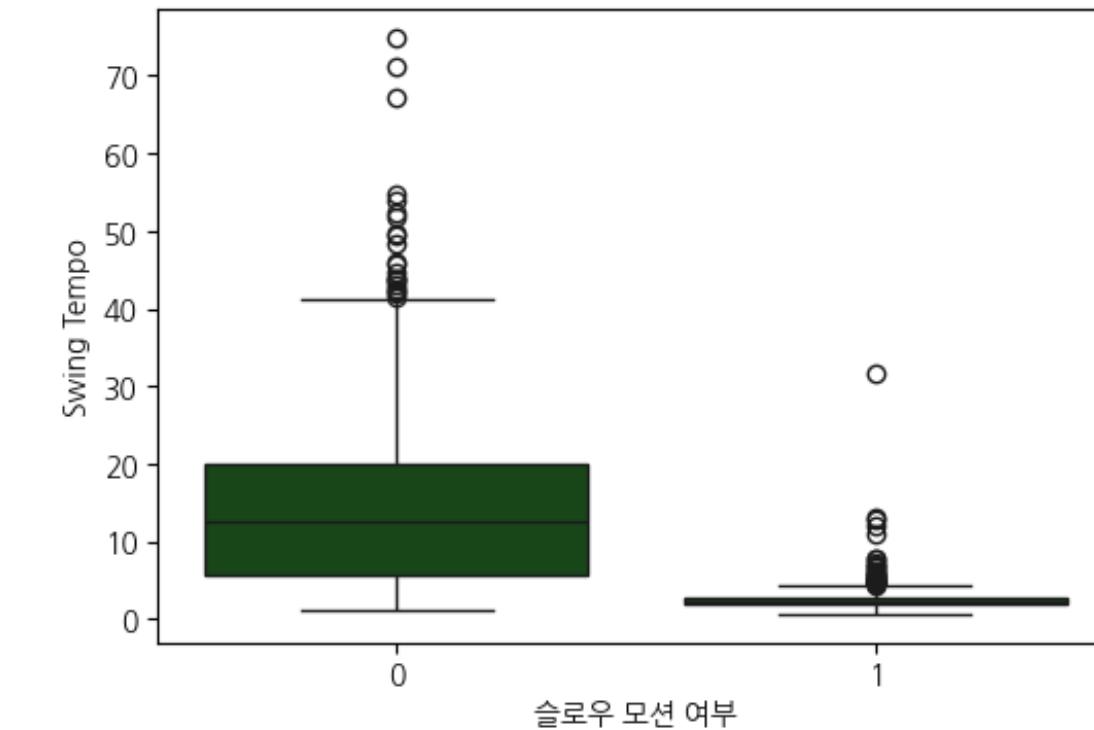
3

[프레임 단위 이벤트 분포]



4

[실시간 vs 슬로우 Swing Tempo]



[swing_tempo = Back-swing 프레임 \div Down-swing 프레임]

- 초반(Address · Toe-Up) 프레임이 70 % 이상
⇒ 모델이 “스윙 시작 자세”에 과도하게 최적화될 가능성 ↑
- 후반(Impact 이후) 프레임은 합계 14 % 미만
⇒ Finish 인식 정확도 ↓, 클래스 가중치/오버샘플링 필요

- 중앙값: 실시간 12.5 F | 슬로우 2.3 F
⇒ 슬로우는 프레임 밀집도가 5배 높아 Feature Scaling 필요
- Tempo 분포의 차이가 크므로 SwingNetV2 학습 시
fps 차이를 보정하지 않으면 과적합으로 인해 정확도 ↓

SwingNetV2_annotation



[Youtube 영상 수집]

- 국내 남녀 프로 골퍼 및 아마추어 선수
- 다양한 실내외 환경에서 촬영한 영상
- 3분 미만 화면비 무관한 shorts 영상 위주 수집
- 스윙자세 전과정 원테이크 영상

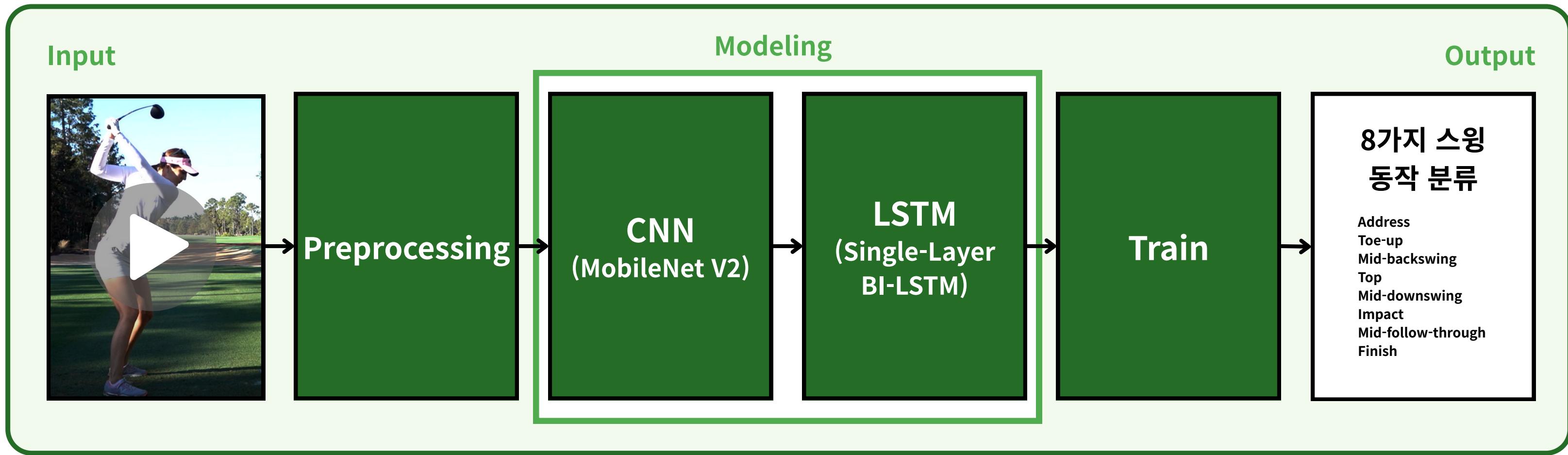
[Annotation Tool : CVAT]

- 기준 자세 8가지 동일
- BBOX 라벨링 후 (0, 1) Nomalization 적용
- 전처리를 위해 일정하게 인물 위치 crop



모델 개발

SwingNetV1



Config.	Input Size (d)	Seq-Length (T)	Batch Size	ImageNet Weights	LSTM Layers (N)	LSTM Hidden (H)	Bidirect.	Params (10 ⁶)	FLOPs (10 ⁹)	PCE (10k iter.)
0	224	32	6	Yes	2	128	Yes	4.07	10.32	66.8
1	224	32	6	No	2	128	Yes	4.07	10.32	1.5
2	224	32	6	Yes	2	128	No	3.08	10.26	54.7
3	192	32	6	Yes	2	128	Yes	4.07	7.62	45.7
4	160	32	6	Yes	2	128	Yes	4.07	5.33	62.4
5	128	32	6	Yes	2	128	Yes	4.07	3.45	57.7
6	160	64	6	Yes	2	128	Yes	4.07	10.65	71.1
7	160	32	12	Yes	2	128	Yes	4.07	5.33	70.1
8	224	32	6	Yes	1	128	Yes	3.67	10.39	69.4
9	224	32	6	Yes	2	64	Yes	3.01	10.26	66.9
10	224	32	6	Yes	2	256	Yes	6.96	10.51	69.3

SwingNetV1_문제점

1. 성능지표
2. CNN Backbone
3. LSTM Layer
4. Augmentation 및 전처리
5. 학습 테크닉
6. 단순 8개 동작 시퀀싱



SwingNetV1_Profiling

*GolfDB논문 성능지표

Seq. Length (T)	FLOPs (10^9)	CPU (ms)*	PCE
64	10.92	10.6	76.2
32	5.41	10.8	74.0
16	2.70	11.5	71.0
8	1.35	12.0	66.0
4	0.68	13.8	63.1

Self CPU time total: 467.273ms
Self CUDA time total: 467.606ms

CPU

Name	Self CPU %	Self CPU	CPU total %	CPU total	CPU time avg	CUDA Mem	Self CUDA Mem	# of Calls
cudaEventRecord	61.94%	289.419ms	61.94%	289.419ms	125.235us	0 b	0 b	2311
cudaDeviceSynchronize	22.61%	105.635ms	22.61%	105.635ms	105.635ms	0 b	0 b	1
cudaLaunchKernel	6.51%	30.400ms	6.51%	30.400ms	60.678us	0 b	0 b	501
CNN_PASS	2.05%	9.563ms	74.82%	349.592ms	349.592ms	5.67 Gb	-15.97 Gb	1
aten::cudnn_batch_norm	0.83%	3.868ms	22.80%	106.524ms	2.049ms	9.75 Gb	0 b	52
aten::_convolution	0.80%	3.727ms	15.61%	72.961ms	1.403ms	9.75 Gb	0 b	52
aten::zeros	0.60%	2.817ms	0.62%	2.909ms	1.455ms	0 b	0 b	2
aten::hardtanh_	0.52%	2.423ms	11.68%	54.580ms	1.559ms	1.81 Gb	0 b	35

 CPU 비효율 개선 필요
[이벤트 기록 + 동기화 + 커널 런칭]
대기시간이 차지하는 비율 : 91%

 GPU 작동 중 병렬로 사용 가능
전처리 과정 CPU 1~2%미만 소모

Self CPU time total: 467.273ms
Self CUDA time total: 467.606ms

GPU(CUDA)

Name	Self CPU %	Self CUDA	Self CUDA %	CUDA total	CUDA time avg	CUDA Mem	Self CUDA Mem	# of Calls
aten::cudnn_batch_norm	0.83%	180.239ms	38.55%	181.112ms	3.483ms	9.75 Gb	0 b	52
aten::cudnn_convolution	0.38%	91.710ms	19.61%	91.710ms	2.620ms	6.39 Gb	6.39 Gb	35
aten::_conv_depthwise2d	0.08%	82.690ms	17.68%	82.758ms	4.868ms	3.36 Gb	0 b	17
aten::clamp	0.13%	78.345ms	16.75%	78.415ms	2.240ms	0 b	0 b	35
aten::copy_	0.06%	16.042ms	3.43%	16.042ms	763.905us	0 b	0 b	21
aten::add	0.04%	4.116ms	0.88%	4.116ms	411.600us	323.44 Mb	323.44 Mb	10
aten::_convolution	0.80%	3.331ms	0.71%	177.799ms	3.419ms	9.75 Gb	0 b	52
aten::cudnn_rnn	0.19%	2.865ms	0.61%	2.943ms	2.943ms	10.02 Mb	53.93 Mb	1

 backbone의 효율/경량화 필요
상위 4개의 cnn연산이 GPU의 92%

 레이어 추가시 정확도 향상 기대
LSTM의 소요 비중 1%미만

SwingNetV1 기반의 새로운 모델

SwingNetV2 제시



사용자의 스윙 영상을
시퀀스별로 분류하여 골프 스윙의 객관적 지표를
구현한 SwingNet 기반의 시스템



Augmentation

시도



Random Horizontal Flip

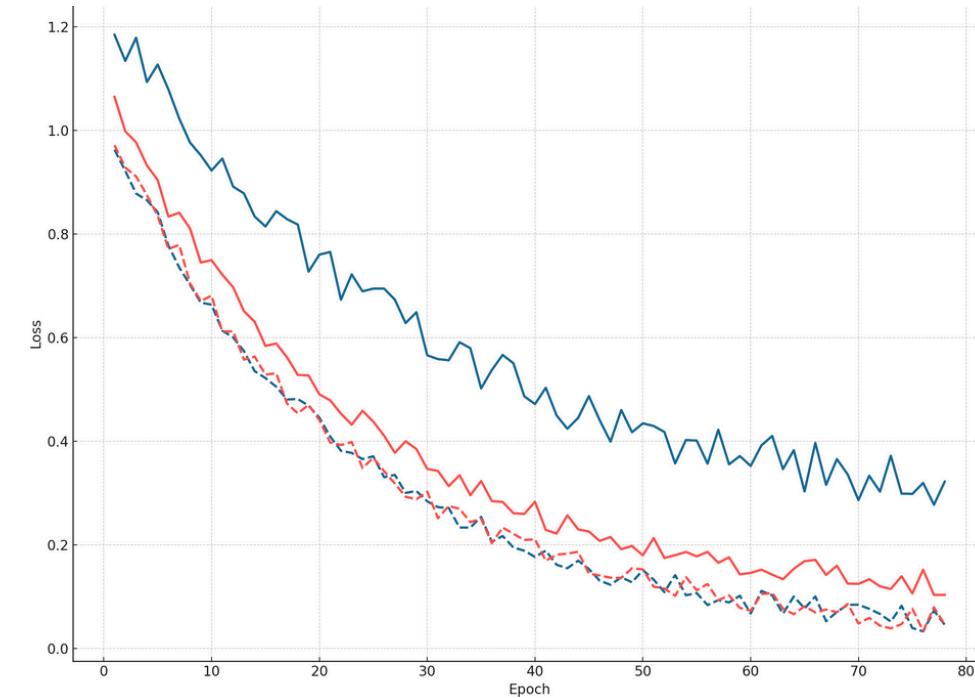
왼손/오른손잡이 고려하여 $p=0.5$ 증강

Random Rotate

-5° ~ 5° 카메라 각도 불균형 보정

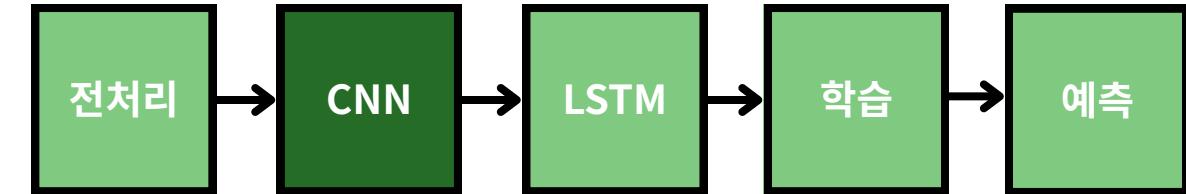
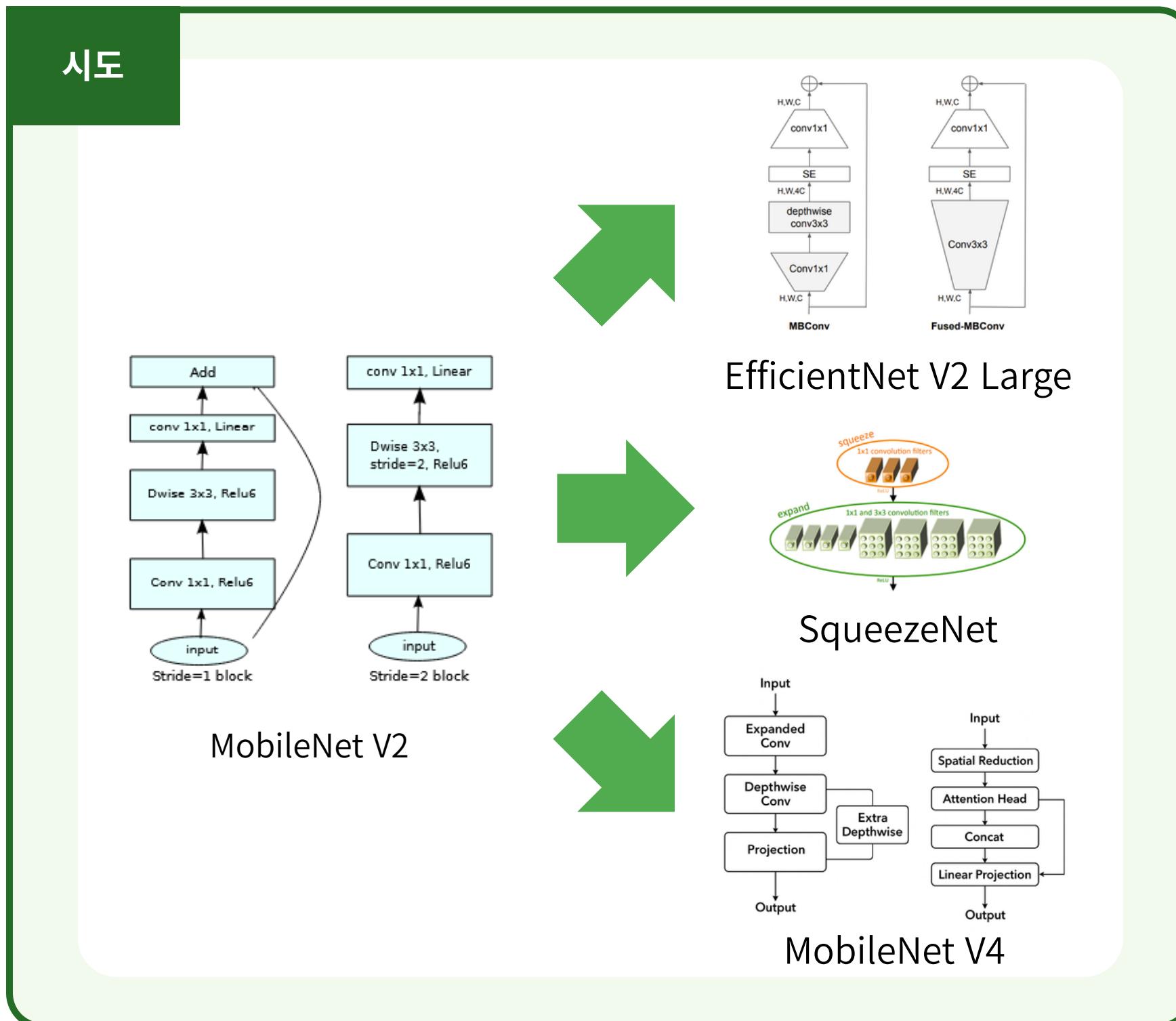
결과

Train Loss (Before Aug)
Val Loss (Before Aug)
Train Loss (After Aug)
Val Loss (After Aug)



Before/After Graph

CNN 변경



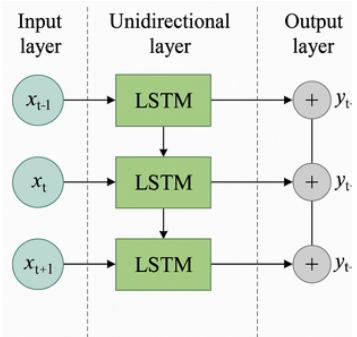
결과 Backbone 성능 비교

모델	Params (M)	FLOPs (G)	Latency (ms)	Top-1 Acc.	PCE (%)
MobileNetV2 (2017)	3.4	0.6	4	71.88%	76.1
SqueezeNet (2016)	1.2	0.72	4.8	58.18%	74.5(E)
Efficient Net V2-B0 (2021)	118.5	56.08	378	85.81%	80.6(E)
Mobilenet V4 (2024)	3.8	0.2	1.3	73.80%	77.8(E)

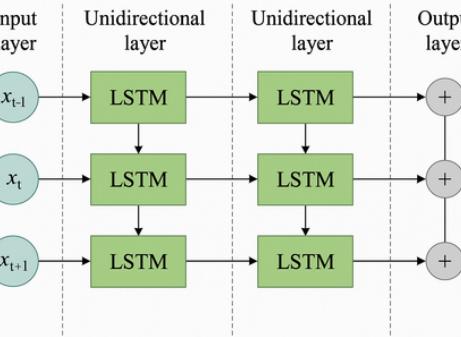
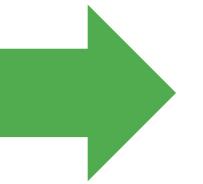
실시간성: MobileNetV4 (1.3 ms) > MobileNetV2(4ms) > SqueezeNet (~4.8 ms)
 정확도: EfficientNet V2 Large > MobileNetV4 > MobileNetV2

적은 데이터 환경에서 Accuracy 대비 Latency 성능
“MobileNetV4”가 가장 우수

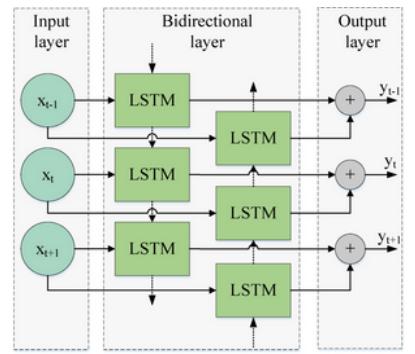
LSTM 변경

시도


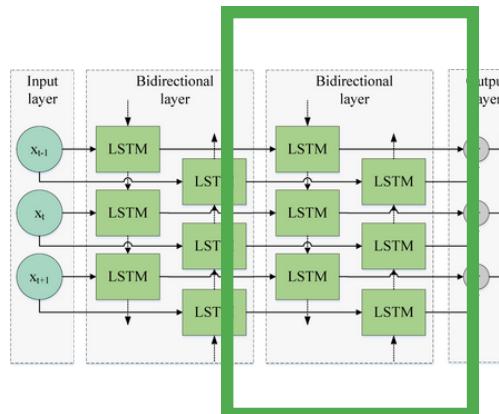
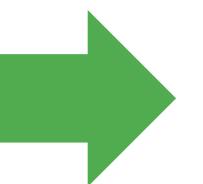
1 Layer
Uni-directional



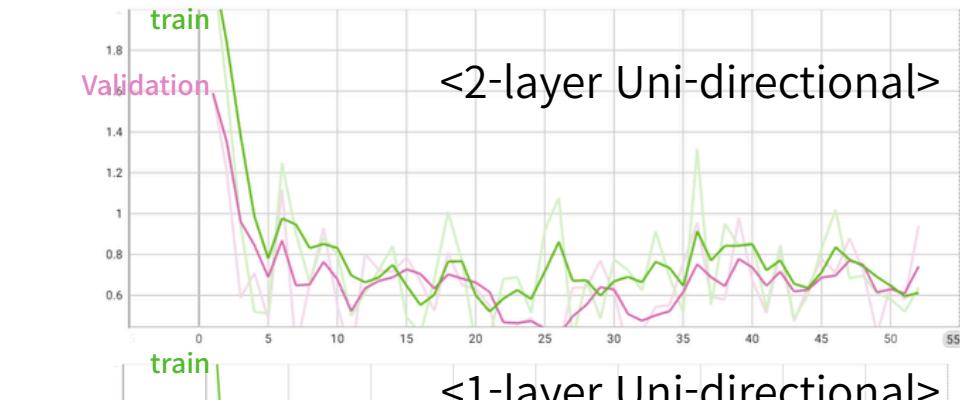
2 Layer
Uni-directional



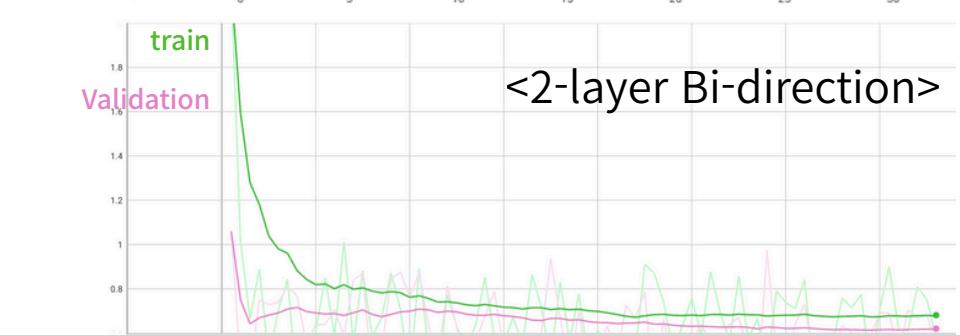
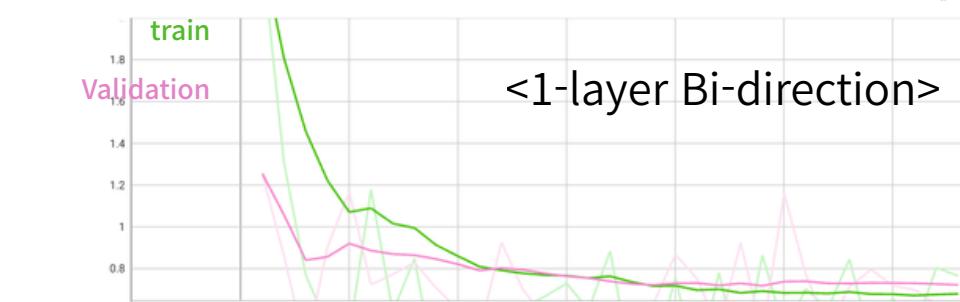
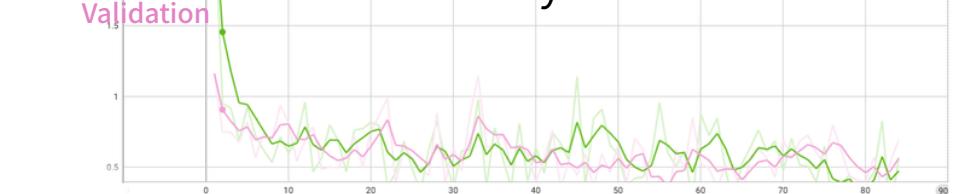
1 Layer
Bi-directional



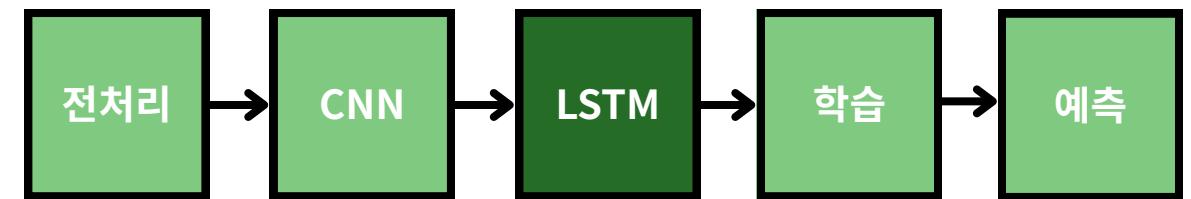
2 Layer
Bi-directional

결과


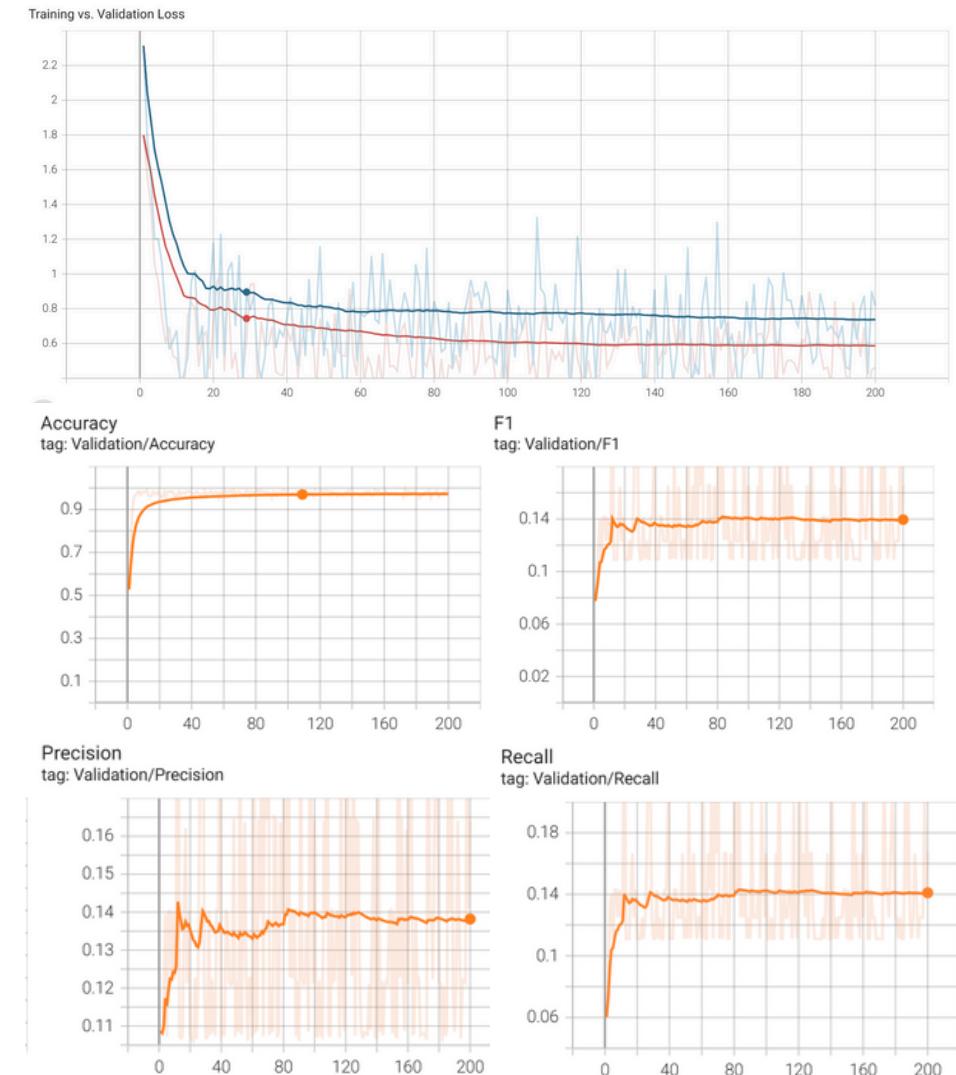
<1-layer Uni-directional>



ConVLSTM



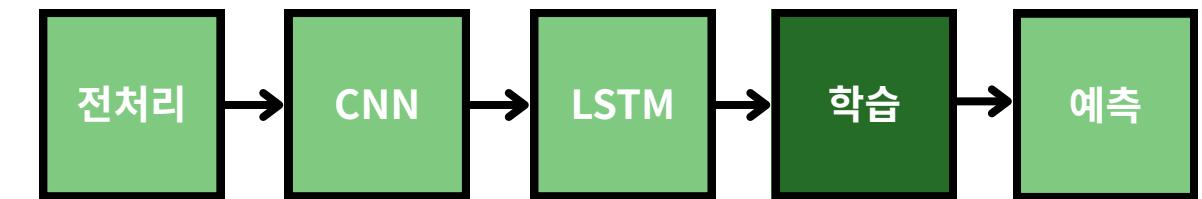
시도



결과

- Training vs Validation Loss**
에폭 0-20: 손실 급락 → 이후 0.5 수준에서 안정화
Training/Validation 차이 작아 과적합 적음
- 정확도(Accuracy)**
에폭 20 이전에 0.90 도달 → 최종 0.92 수준 유지
- F1-Score 및 Precision**
초기 0.11 → 0.14 상승 후 정체 → 클래스 불균형
- Recall**
초기 0.06 → 0.14로 개선 후 0.14 수준 머무름
→ 불균형 영향 지속

학습 Hyper-Parameter Tunning



시도

bs = batch size
early-stopping 적용
gradient accumulation 적용

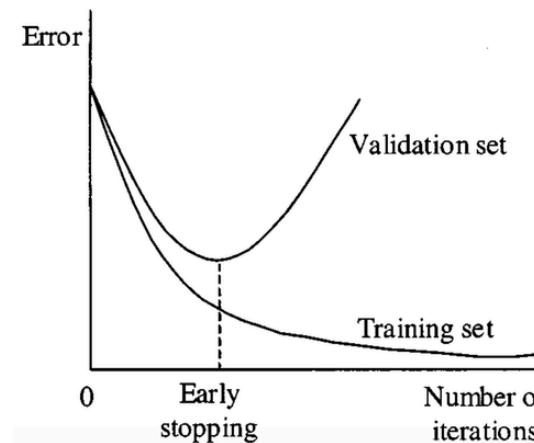
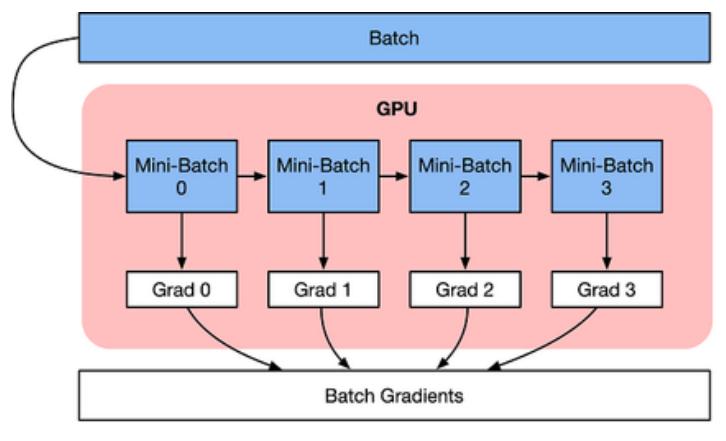
Loss Function function

$$Loss(\theta) = \frac{1}{2} (h(x^k) - y^k)^2$$

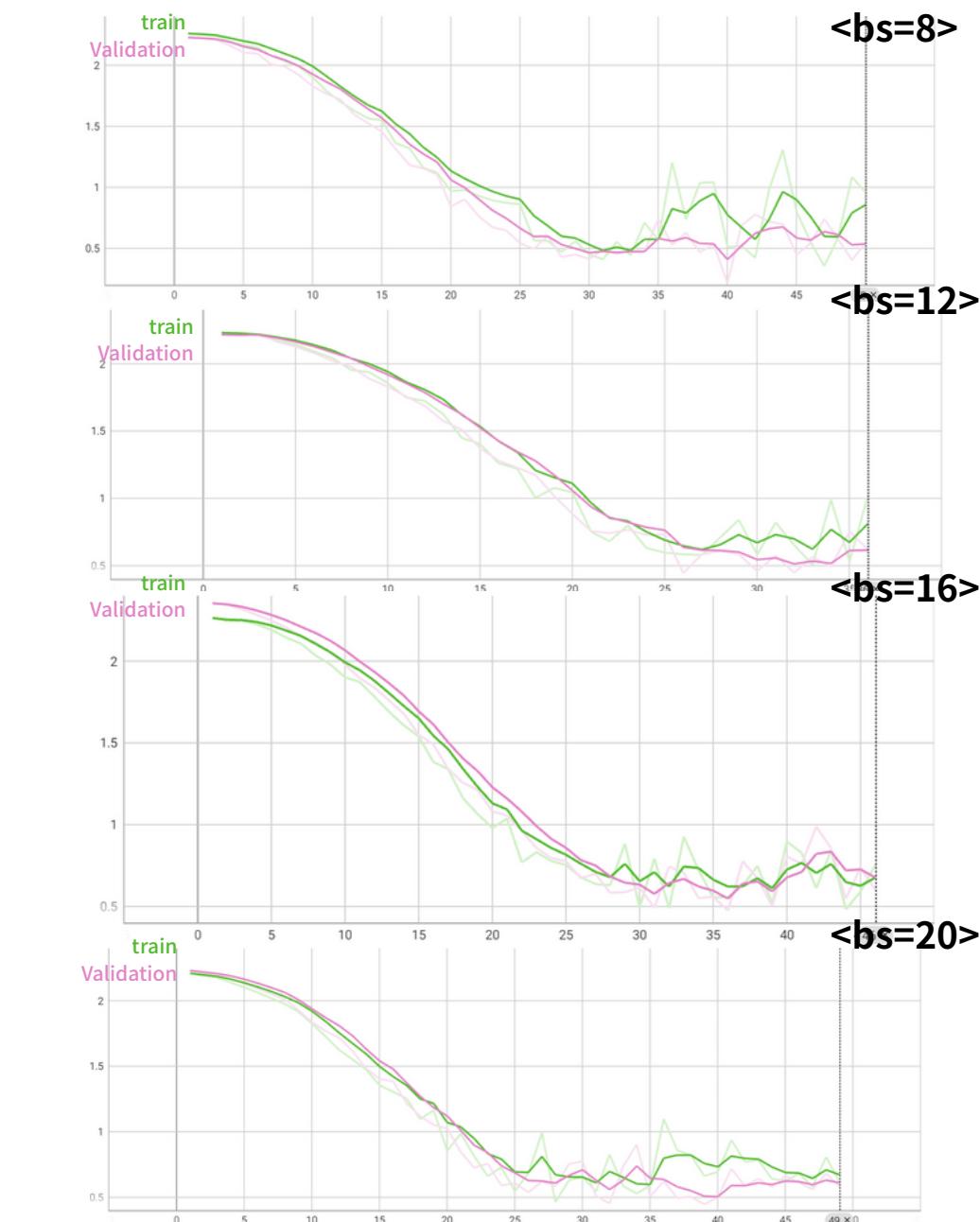
Gradient and weight function

$$accumulated = \sum_{i=0}^N grad_i$$

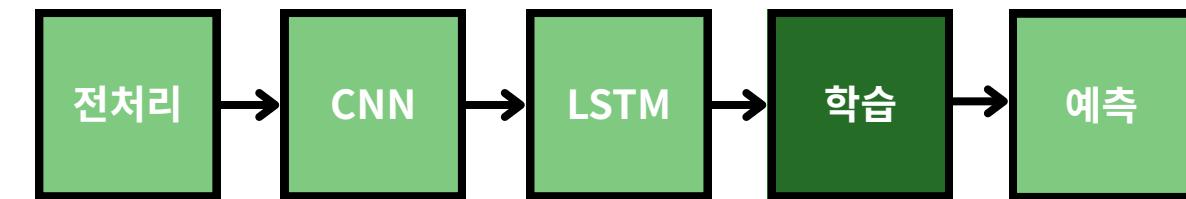
$$\theta_i = \theta_{i-1} - lr * \sum_{i=0}^N grad_i$$



결과



Learning Rate Scheduler / Warm up



시도

LAMBDA LR

$$lr_{epoch} = lr_{initial} * \text{Lambda}(epoch)$$

$$lrepoch = lr_{initial} * \text{Lambda}(epoch)$$

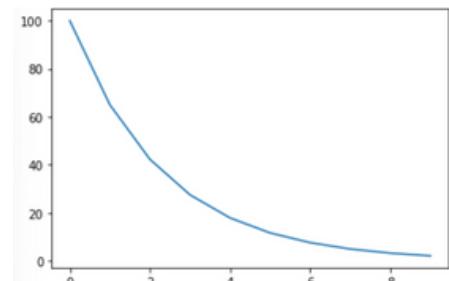
ExponentialLR

$$lr_{epoch} = \text{Gamma} * lr_{epoch - 1}$$

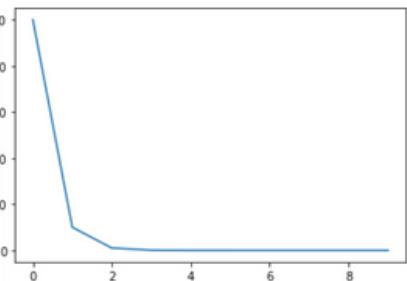
$$lrepoch = \text{Gamma} * lr_{epoch - 1}$$

CosineAnnealingLR

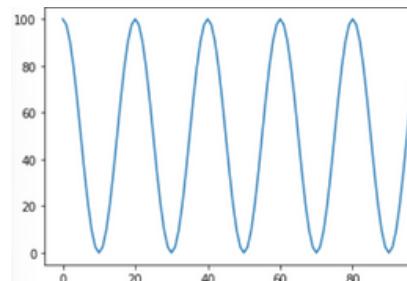
$$\eta_t = \eta_{min} + \frac{1}{2}(\eta_{max} - \eta_{min}) \left(1 + \cos \left(\frac{T_{cur}}{T_{max}} \pi \right) \right)$$



LAMBDA LR

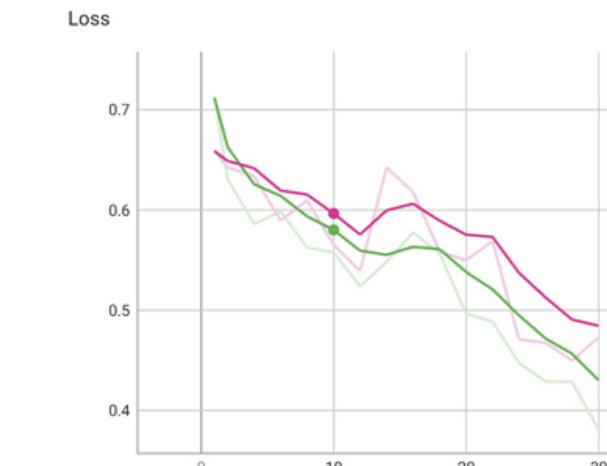


ExponentialLR



CosineAnnealingLR

결과

**Warm-up:** 초기 안정화를 달성**Scheduler:** 후반 성능 극대화

가이드 활용 지표

세 점 사이의 각도 계산 - angle

$$r = \text{atan}2(y_C - y_B, x_C - x_B) - \text{atan}2(y_A - y_B, x_A - x_B),$$

$$\alpha = \left| \frac{180}{\pi} r \right|,$$

$$\text{angle} = \begin{cases} 360 - \alpha, & \alpha > 180, \\ \alpha, & \alpha \leq 180. \end{cases}$$

두 점 사이의 기울기 계산 - inclination

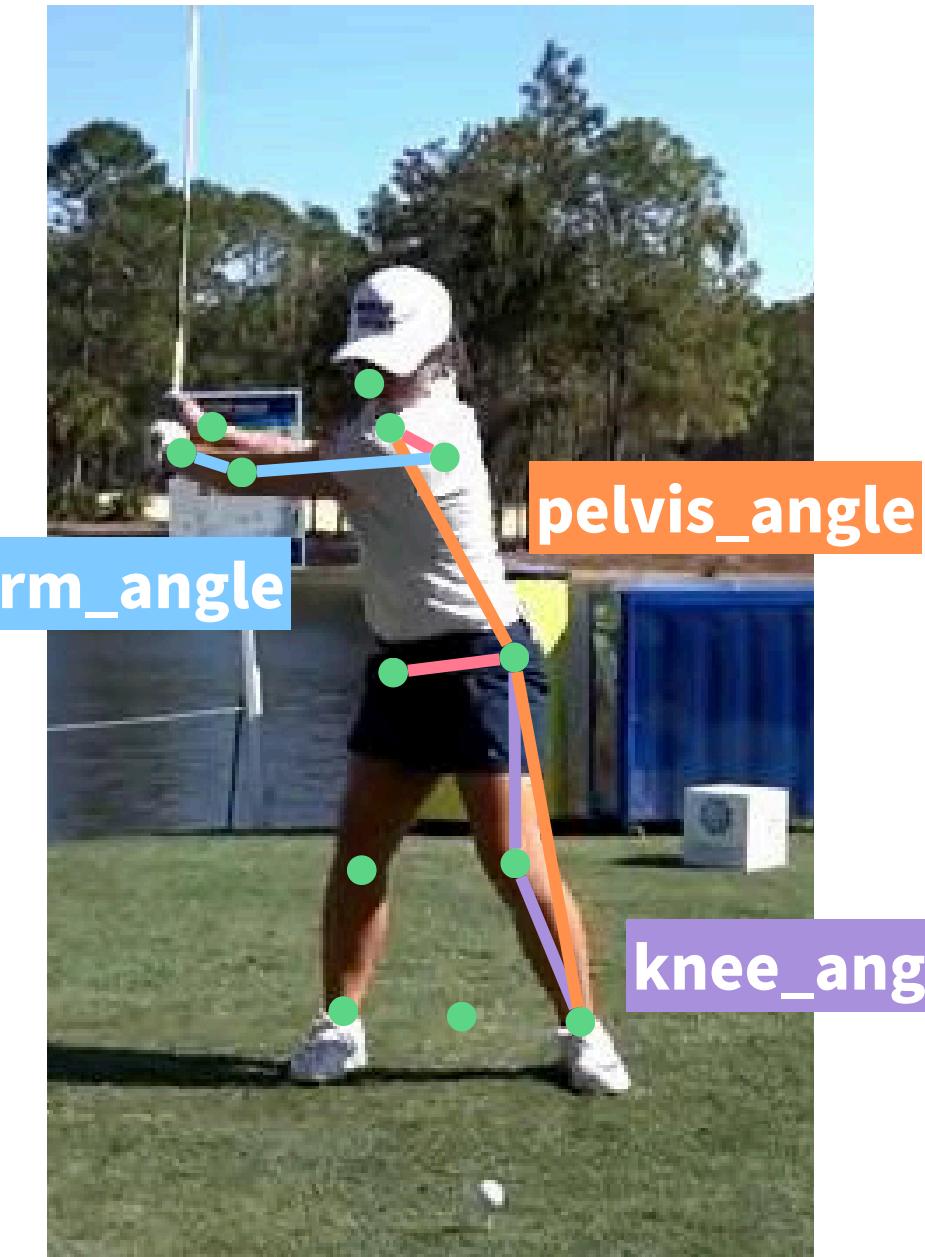
$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2},$$

$$d \cdot x_1 = 0 \implies \boxed{\text{angle} = 0},$$

$$\cos \theta = \begin{cases} \frac{(x_2 - x_1)(-x_1)}{d x_1}, & \text{axis} = x, \\ \frac{(y_2 - y_1)(-y_1)}{d y_1}, & \text{axis} = y, \end{cases} \quad K = \text{int}\left(\frac{180}{\pi}\right),$$

$$\theta = \arccos(\cos \theta),$$

$$\text{angle} = \begin{cases} K \theta, & \text{orientation} = \text{right}, \\ 180 - K \theta, & \text{orientation} = \text{left}. \end{cases}$$



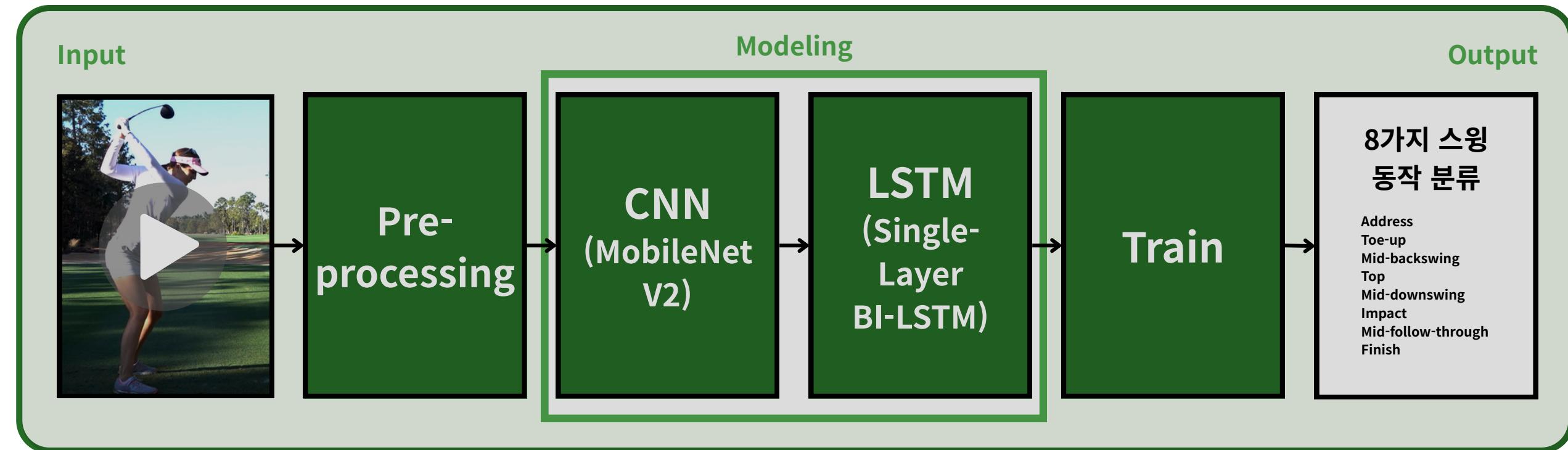
point : 13

angle : 3

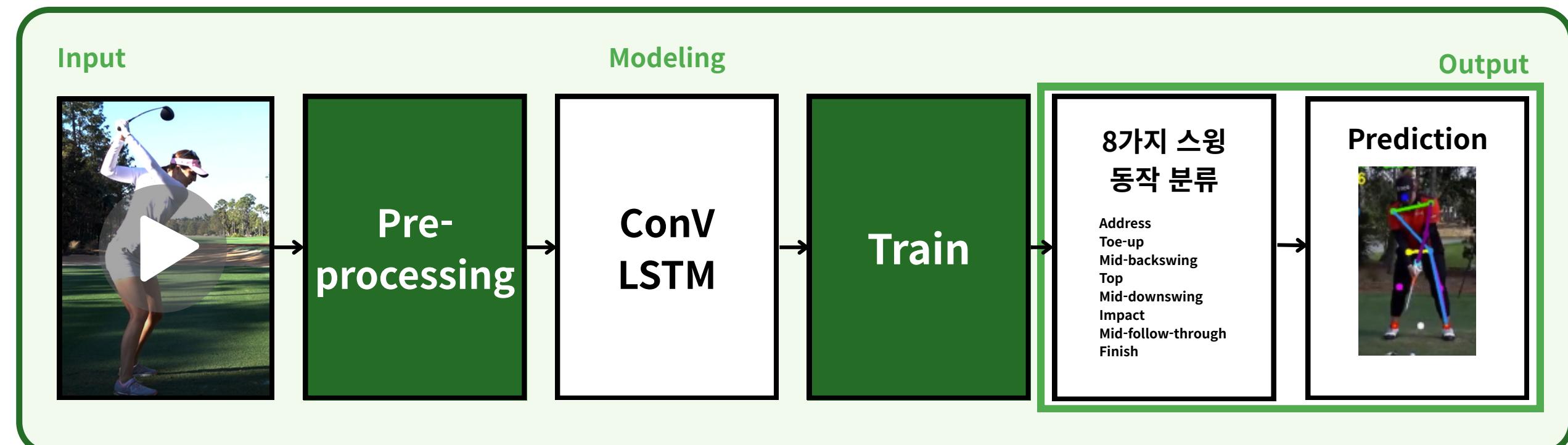
inclination : 2

SwingNetV1, V2 비교

SwingNetV1



SwingNetV2





검증 및 테스트

검증 및 테스트

- input video : <https://www.youtube.com/watch?v=3wru0WH0buk>



Address



Toe-up



Mid-backswing



Top



Mid-downswing



Impact



Mid-follow-through

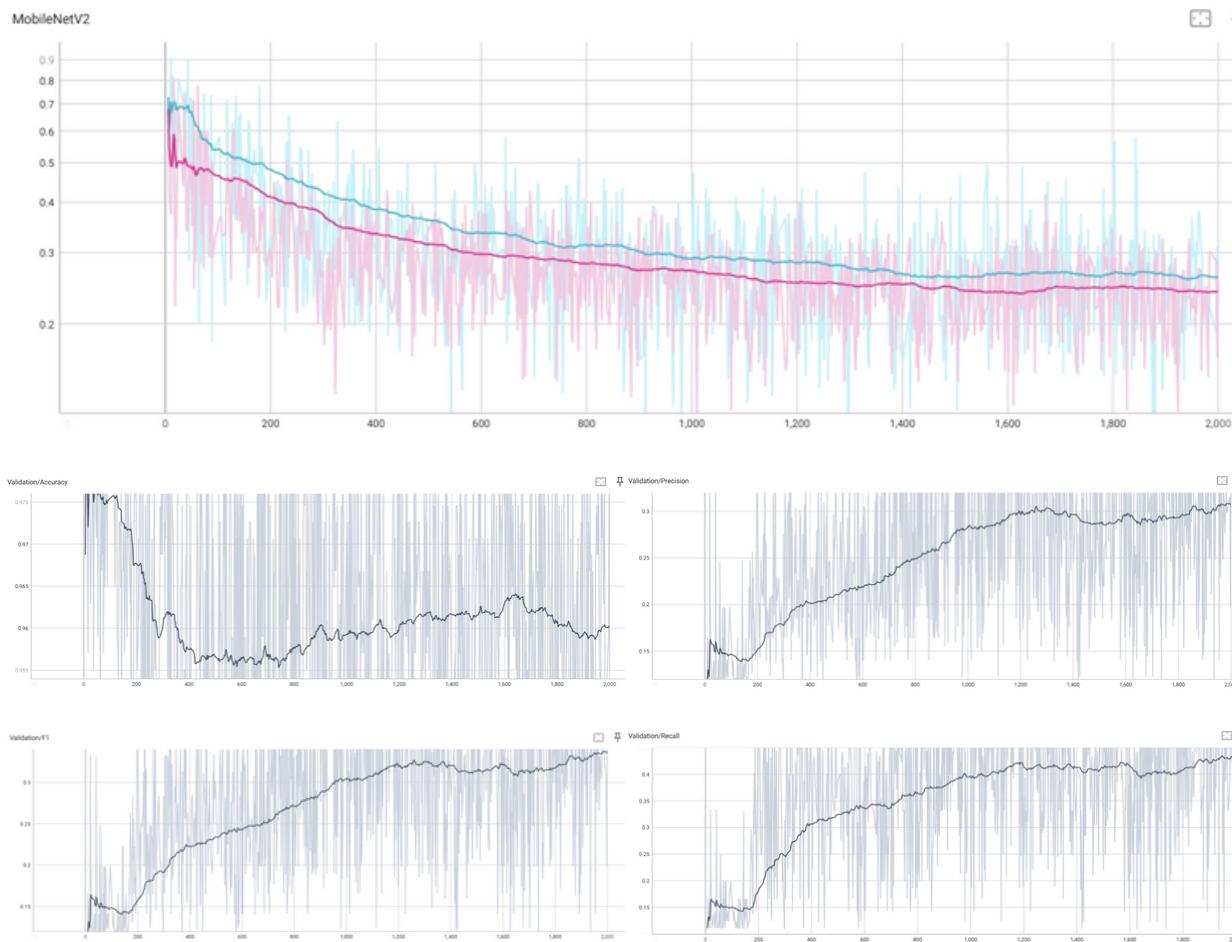


Finish

검증 및 테스트_2

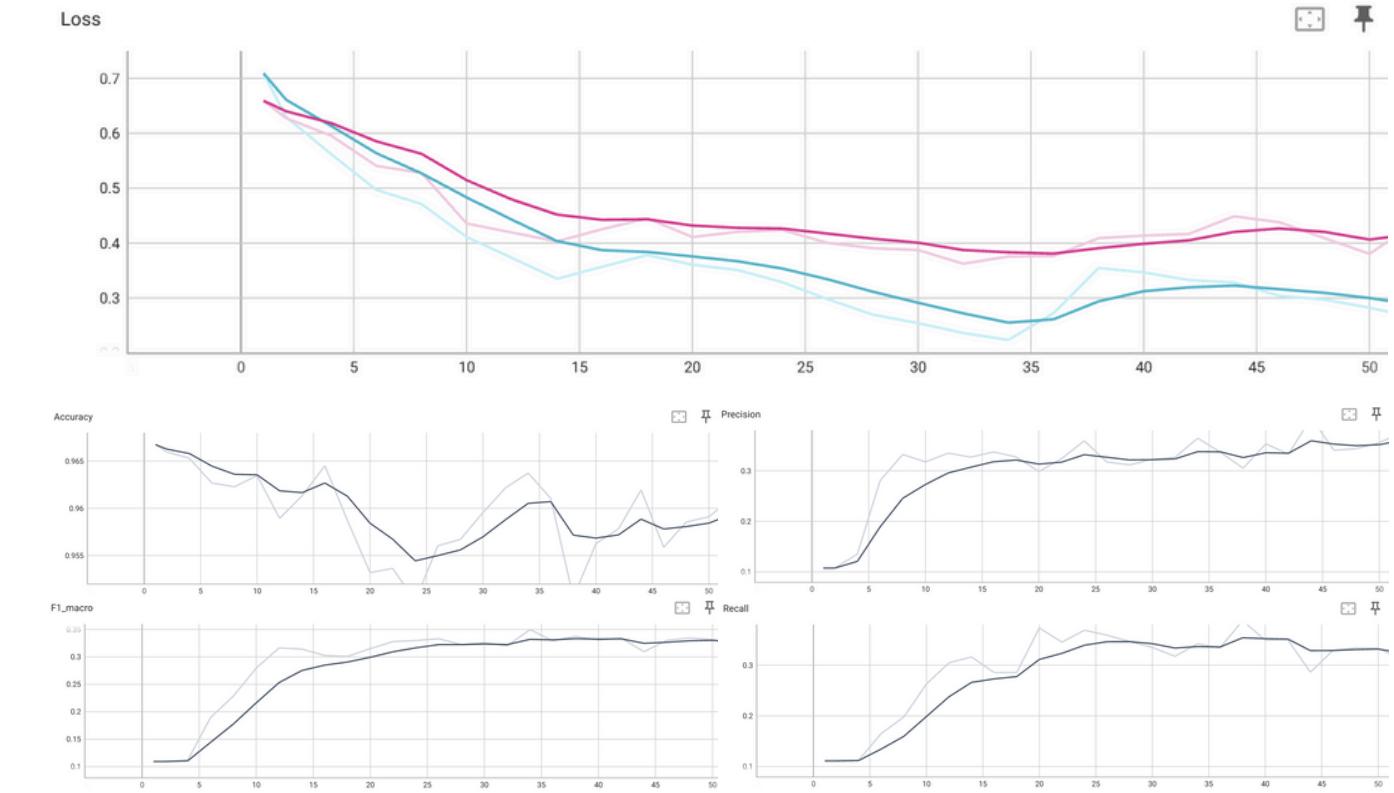
V1

SwingNetV1



V2

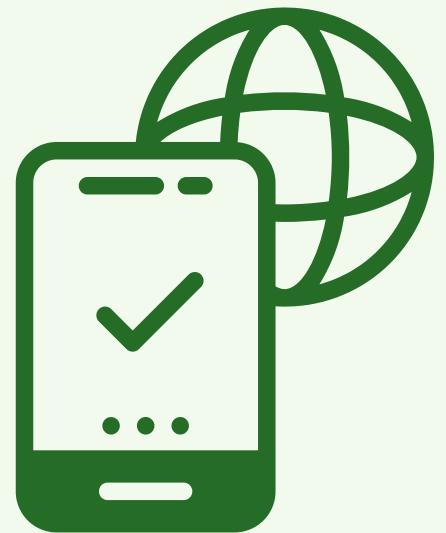
SwingNetV2





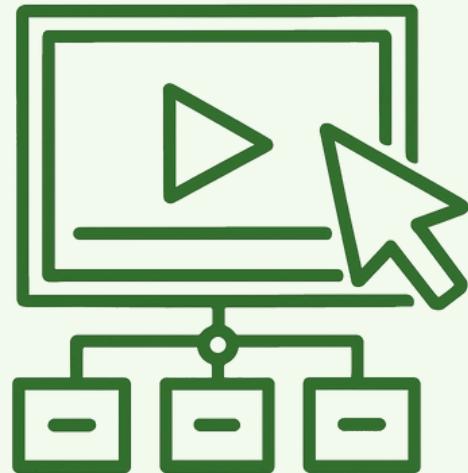
기대효과

기대효과_1 사용자 관점



접근성 제고

기존 보유한 영상으로
손쉽게 이용 가능



코칭용 클립 자동 생성

대표 프레임 추출
→ 단계별 피드백 자료 확보



개인별 스윙 통계 리포트

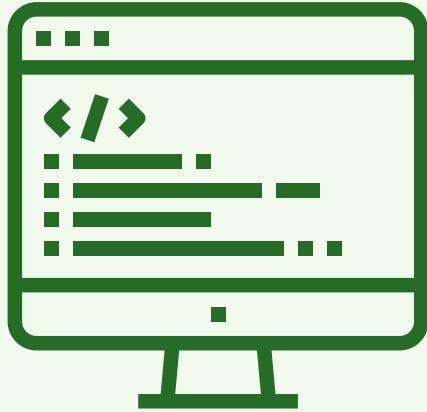
누적된 이벤트 분류를 분석
→ 반복 오류 패턴 및 개선 추이

기대효과_2 비즈니스 관점



데이터 자산화

라벨링된 시퀀스 DB 구축·활용



라벨링 자동화

골프영상 시퀀스 자동 라벨링
→ 운영 비용 ↓



스윙자세 피드백 서비스로 확대

HPE모델과 결합하여
기존 대비 정교한

스윙자세 피드백 서비스로 확대 기대

감사합니다

1조 서상원 안지현 임연오

SwingNetV2



1조 서상원 안지현 임연오
2025.06.19

