

해상도 및 노이즈 변화에 따른 Mediapipe BlazePose의 이미지 기반 포즈 추정 성능 분석

이건아¹, 한요섭^{1,2}

숭실대학교 전자정보공학부¹

숭실대학교 지능형반도체학과²

kuna001@soongsil.ac.kr, yoseob.han@ssu.ac.kr

Mediapipe BlazePose with Resolution and Noise Changes Image-based pose estimation performance analysis

Geona Lee¹, Yoseob Han^{1,2}

¹Department of Electronic Engineering, Soongsil University

²Department of Intelligent Semiconductors, Soongsil University

요약

본 논문은 Google Mediapipe에서 제공하는 BlazePose 모델을 기반으로, 다양한 해상도 및 이미지 노이즈 환경에서의 포즈 추정 성능을 LSP(Leeds Sports Pose) 데이터셋을 활용하여 비교 분석하였다. 기존의 연구들은 일반적으로 고해상도 이미지에서의 Pose Estimation 성능을 보고하였으나 BlazePose는 모바일 환경을 겨냥한 경량 모델로 설계되어 고해상도 및 노이즈가 포함된 이미지에서의 성능 저하가 발생할 수 있다. 이에 따라 본 논문에서는 네 가지의 해상도 및 세 가지 종류의 노이즈가 BlazePose 성능에 미치는 영향을 분석하였으며, PCK(Percentage of Correct Keypoints)를 주요 평가 지표로 활용하였다. 또한, 이미지 전처리 순서가 결과에 미치는 영향을 분석하기 위해 Gaussian 노이즈를 활용하여 Denoising 후 Resize와 Resize 후 Denoising 두 방식 간의 성능 차이도 비교 실현하였다.

I. 서론

포즈 추정(Pose Estimation)은 영상에서 사람의 관절 위치를 추정하는 기술로, 스포츠 분석, 헬스케어, HCI(Human-Computer Interaction) 등 다양한 분야에 활용되고 있다. 최근에는 모바일 및 웹캠 환경에서도 실시간으로 작동할 수 있는 경량화된 포즈 추정 모델에 대한 수요가 증가하고 있으며 Google Mediapipe의 BlazePose는 이러한 요구를 충족하기 위해 설계된 대표적인 lightweight 모델로 저해상도에서도 빠르고 안정적인 추론이 가능하다.

기존의 연구 대부분은 모델이 실질적으로 어떤 해상도 조건에서 가장 적합하게 동작하는지에 대한 분석은 부족하다. 특히 BlazePose는 고해상도 이미지 입력 시 내부 downsampling 구조의 비효율성으로 인해 오히려 성능이 하락하는 현상이 나타날 수 있으며 이에 대한 정량적 실험이 필요하다.

본 논문에서는 고정된 구조의 BlazePose 모델을 대상으로 다양한 해상도와 노이즈를 적용한 상태에서의 성능 변화를 분석하였다. 평가에는 LSP(Leeds Sports

Pose) 데이터셋을 사용하였으며 PCK(Percentage of Correct Keypoints) 지표를 기준으로 정확도를 분석하였다. 또한 이미지 전처리 순서가 결과에 미치는 영향을 확인하기 위해 Denoising 후 Resize 와 Resize 후 Denoising 의 두 절차를 각각 실험하였다.

II. 본론

1. 실험환경

모델	Mediapipe BlazePose
데이터셋	LSP(Leeds Sports Pose)
해상도	320x240, 640x480, 1280x720, 1920x1080
노이즈	Gaussian, Salt & Pepper, Uniform
전처리 순서	Denoising → Resize, Resize → Denoising
평가지표	PCK@30px

표 1: 실험 환경

2. 실험

2.1. 해상도 변화 실험

해상도 변화에 따른 BlazePose의 성능을 분석하기 위해 동일한 LSP 이미지에 대해 네 가지 해상도(320x240, 640x480, 1280x720, 1920x1080)로 Resize한 후 포즈 추정을 수행한 결과를 표 2에 정리하였다.

해상도	PCK(%)
320x240	68.49
640x480	48.50
1280x720	15.30
1920x1080	7.90

표 2: 해상도에 따른 PCK 평가 결과(%)

320x240에서의 정확도가 고해상도의 성능보다 높은 것을 확인할 수 있었고 고해상도에서는 성능이 급격하게 하락하였다. 이는 BlazePose 내부에 존재하는 downsampling 기반 전처리 구조가 고해상도 입력 이미지의 세부 정보를 압축하거나 내부 해상도 정규화 과정에서 불안정한 예측을 초래한 것으로 해석할 수 있다.

2.2. 해상도 및 노이즈 동시 적용 실험

모든 해상도마다 노이즈를 적용한 12가지 경우에 대해 동일하게 BlazePose를 적용한 결과를 표3에 정리하였다.

해상도	Gaussian	Salt & Pepper	Uniform
Gaussian	67.13	47.58	63.03
640x480	48.43	34.06	47.24
1280x720	8.39	5.58	8.85
1920x1080	4.24	2.54	3.85

표 3: 각 해상도 및 노이즈에 따른 PCK 평가 결과(%)

해상도가 낮은 경우에는 Gaussian 및 Uniform 노이즈가 성능에 상대적으로 적은 영향을 미친 반면, Salt & Pepper 노이즈는 전 해상도에서 가장 큰 성능 저하를 보였다. 이는 Salt & Pepper 노이즈가 꽉 채워진 단위의 강한 이상치를 포함하여 모델의 관절 예측 정확도에 치명적인 간섭을 일으키기 때문이다. 또한 고해상도에서는 원본 이미지 자체의 성능이 낮게 때문에 노이즈로 인한 추가적인 성능 저하는 상대적으로 적게 나타났으며, BlazePose가 고해상도에 취약한 구조임을 확인하였다.

2.3. 전처리 순서 비교 실험

동일한 Gaussian 노이즈가 적용된 이미지에 대해 Denoising과 Resize의 순서를 바꾸어 실험한 결과를 표 4에 정리하였다.

전처리 방식	PCK(%)
Denoising → Resize	45.26

Resize → Denoising	45.38
--------------------	-------

표 4: 전처리 순서에 따른 PCK 평가 결과(%)

결과적으로 두 방식 간의 성능 차이는 0.12% 정도로 미미하게 나타났으며 Gaussian 노이즈에 대해서는 전처리 순서가 BlazePose 성능에 유의미한 영향을 미치지 않음을 확인할 수 있었다.

III. 결론

본 논문에서는 BlazePose 모델을 대상으로 다양한 해상도 및 노이즈 환경에서의 포즈 추정 성능을 LSP 데이터셋을 기반으로 분석하였다. 실험 결과, 중저해상도(320×240, 640×480)에서 가장 높은 정확도를 보였으며, 고해상도로 갈수록 오히려 성능이 저하되는 현상이 확인되었다. 이는 BlazePose의 downsampling 기반 구조가 고해상도 입력에 적합하지 않음을 시사한다.

또한, Gaussian 및 Uniform 노이즈는 상대적으로 안정적인 성능을 보인 반면, Salt & Pepper 노이즈는 모든 해상도에서 큰 성능 저하를 유발하였다. 마지막으로, 전처리 순서를 달리한 실험에서는 Denoising과 Resize 순서에 따른 성능 차이가 거의 없으며, BlazePose는 Gaussian 노이즈에 대해 전처리 순서에 크게 영향을 받지 않음을 확인하였다.

이러한 결과는 BlazePose와 같은 경량 포즈 추정 모델을 실환경에 적용할 때 해상도 및 노이즈 처리 전략이 중요한 요소임을 시사한다.

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원-대학 ICT 연구센터(ITRC)의 지원을 받아 수행된 연구임(IITP-2025-RS-2020-II201602)

참 고 문 헌

- [1] Bazarevsky, V., Kartynnik, Y., Vakunov, A., Tkachenka, A., & Grundmann, M. (2020). BlazePose: On-device real-time body pose tracking. arXiv preprint arXiv:2006.10204.
- [2] Andriluka, M., Pishchulin, L., Gehler, P., & Schiele, B. (2014). 2D human pose estimation: New benchmark and state of the art analysis. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3686–3693).