

생성형 AI 인물 이미지의 안면 정체성 유지와 포즈 제어를 위한 하이브리드 워크플로우 연구

Krea 2 Identity·ReID 비교와 ComfyUI 기반 기능적 참조 분리를 중심으로

A Study on a Hybrid Workflow for Facial Identity Preservation and Pose Control in AI-Generated Human Images

Focusing on the Comparison of Krea 2 Identity and ReID and the Functional Separation of References in ComfyUI

주 저 자 : 한규훈 (Han, Gyu Hoon) 광운대학교 미디어커뮤니케이션학부 겸임교수

교 신 저 자 : 오문석 (Oh, Moon Seok) 광운대학교 미디어커뮤니케이션학부 교수
motion@kw.ac.kr

<https://doi.org/10.46248/kidrs.2026.3.680>

접수일 2026. 08. 20. / 심사완료일 2026. 08. 28. / 게재확정일 2026. 09. 07. / 게재일 2026. 09. 30.

본 논문은 2026학년도 광운대학교 교내학술연구비를 지원받아 수행되었음.

Abstract

This study compares the facial identity preservation and pose control characteristics of Krea 2-based Identity and ReID methods and proposes a ComfyUI-based hybrid workflow that functionally separates identity and pose references. A total of 135 images were generated using five synthetic identities, three levels of pose variation, and three generation methods. Performance was evaluated using Facial Identity Similarity (FIS), PCK@0.2-based Pose Accuracy (PA), and Inter-Image Identity Consistency (IIC). Identity showed relatively strong facial identity preservation, while ReID performed better in pose reproduction. The Hybrid method achieved an FIS of 0.87, PA of 0.91, and IIC of 0.85, indicating comparatively balanced performance. However, because it includes additional pose conditioning and a two-stage generation structure, the results should be interpreted as an exploratory workflow-level comparison under limited conditions rather than as the independent effect of functional reference separation.

Keyword

ComfyUI, Krea 2, workflow(워크플로우), facial identity(안면 정체성), pose control(포즈 제어)

요약

본 연구는 Krea 2 기반 Identity와 ReID의 안면 정체성 유지 및 포즈 제어 특성을 비교하고, 정체성 참조와 포즈 참조의 역할을 기능적으로 분리한 ComfyUI 기반 하이브리드 워크플로우를 제안하였다. 실험은 5개의 합성 인물과 3단계 포즈 변화 수준을 대상으로 Identity, ReID, 하이브리드 방식에서 총 135장의 이미지를 생성하여 수행하였다. 평가는 안면 정체성 유사도(FIS), PCK@0.2 기반 포즈 정확도(PA), 이미지 간 정체성 일관성(IIC)을 활용하였다. 분석 결과, Identity는 안면 정체성 유지에, ReID는 포즈 재현에 상대적으로 높은 특성을 보였으며, 하이브리드는 FIS 0.87, PA 0.91, IIC 0.85로 세 지표에서 비교적 균형적인 값을 나타냈다. 다만 하이브리드는 별도의 포즈 조건화와 2단계 생성구조를 포함하므로, 이 결과는 기능적 분리의 독립적 효과가 아닌 제한된 조건에서의 탐색적 워크플로우 비교로 해석할 필요가 있다.

목차

1. 서론

- 1-1. 연구의 배경
- 1-2. 연구 목적 및 방법

2. 이론적 배경

- 2-1. 생성형 AI 이미지의 인물 정체성과 포즈 제어
- 2-2. Krea 2 Identity와 ReID

3. ComfyUI 기반 하이브리드 워크플로우 설계

- 3-1. Identity 및 ReID 워크플로우 분석
- 3-2. 정체성 및 포즈 정보 분리 기반 하이브리드 워크플로우 설계
- 3-3. 실험 설계 및 평가방법

4. 실험 결과 및 분석

- 4-1. 생성 방식에 따른 안면 정체성 및 포즈 제어 비교
- 4-2. 포즈 변화 수준에 따른 결과

5. 결론

참고문헌

1. 서론

1-1. 연구의 배경

확산 기반 생성모델은 고품질 이미지 합성과 다양한 조건 제어 능력을 빠르게 발전시키며, 이미지 생성 기술의 활용 범위를 확장하고 있다.¹⁾²⁾ 이러한 흐름 속에서 Krea AI는 2026년 5월 12일 자체적으로 개발한 파운데이션 이미지 모델인 Krea 2를 공개하였다. Krea 2는 텍스트와 이미지 참조를 활용하여 사용자가 의도한 시각적 스타일과 생성 방향을 세밀하게 제어할 수 있도록 설계된 이미지 생성 모델이다. 특히 다양한 미학적 표현과 참조 이미지 기반의 제어 기능을 제공한다.³⁾ 그러나 확산 기반 생성모델은 확률적 샘플링 과정으로 인해 동일하거나 유사한 입력 조건에서도 생성 결과에 변동성이 나타날 수 있으며, 인물 이미지에서는 얼굴형과 이목구비 등 정체성을 구성하는 특징이 일관되게 유지되지 않는 문제가 발생할 수 있다. 특히 동일 인물을 새로운 자세로 재생성하는 과정에서는 인물의 정체성 유지와 목표 포즈 구현이라는 두 가지 조건이 상호 충돌할 수 있다. 참조 인물의 안면 특징을 강하게 보존할 경우, 목표 포즈로의 변화가 제한될 수 있으며, 반대로 포즈의 자유도를 높일 경우, 얼굴형이나 이목구비가 변화하여 동일 인물로 인식되지 않는 현상이 나타날 수 있다. 이러한 문제는 동일한 인물이 반복적으로 등장해야 하는 캐릭터 시트, 광고 시안, 스토리보드 등의 콘텐츠 제작 과정에서 결과물의 일관성과 활용 가능성을 저해하는 핵심적인 품질 제약으로 작용한다.

본 연구에서 활용한 ComfyUI는 모델 로딩, 참조 이미지 입력, 조건화, 샘플링 및 출력 과정을 노드 기반의 그래프로 구성할 수 있는 생성형 AI 인터페이스로, 서로 다른 조건 정보의 역할을 명시적으로 분리하고 동일한 실험 조건을 반복적으로 적용하는 데 적합하다.⁴⁾ 이에 본 연구에서는 ComfyUI를 활용하여 정체성

정보와 포즈 정보의 역할을 기능적으로 분리하고 이에 따른 생성 결과의 성능 특성을 정량적으로 비교·분석하고자 했다. 또한 연구의 범위를 배경이나 의상 등을 포함하는 포괄적인 외형 일관성에 두지 않고, 동일 인물로 인식되기 위한 안면 정체성의 유지와 목표 신체 자세의 구현 정확성에 한정하여 분석하고자 하였다.

1-2. 연구 목적 및 방법

본 연구의 목적은 첫째, Identity와 ReID의 안면 정체성 유지 및 포즈 제어 특성을 비교하고, 둘째, 정체성 참조(WHO)와 포즈 참조(HOW)를 기능적으로 분리한 하이브리드 워크플로우를 설계하며, 셋째, 동일한 인물·포즈·생성 조건에서 세 방식의 성능을 정량적으로 비교·분석하는 데 있다. 특히 본 연구에서 제안하는 하이브리드 워크플로우가 모든 지표에서 절대적으로 우수하다고 전제하기보다, 안면 정체성 유지와 포즈 제어 사이에서 발생하는 상충관계가 하이브리드 방식에서 어떠한 양상으로 나타나는지를 검토하고자 하였다. 이러한 목적을 달성하기 위해 본 연구는 ComfyUI Desktop 환경에서 Krea 2 기반의 Identity, ReID 및 하이브리드 방식으로 각각 이미지 생성 워크플로우를 구성하고, 각 방식의 고유한 입력 구조를 유지하되, 동일한 기준 인물과 목표 포즈 조건을 설정하여 생성 결과를 비교·분석하였다. 실험 과정에서는 기본 모델, 이미지 해상도, 샘플링 조건 및 시드 등 주요 생성 변수를 동일하게 통제하고, 프롬프트는 각 방식의 입력 구조를 유지하되 의미적 범위와 목표 조건을 통일하였다. 또한 포즈 변화 수준을 단계적으로 구분하여 각 방식의 성능 변화를 비교·분석하였다. 다만 하이브리드 방식은 별도의 포즈 조건화와 2단계 생성구조를 포함하므로, 본 연구는 기능적 참조 분리의 독립적 효과를 검증하기보다 세 생성 워크플로우의 상대적 성능 특성을 탐색적으로 비교하는 데 목적을 둔다.

구체적인 연구방법은 다음과 같다. 첫째, Identity와 ReID의 조건화 구조와 생성 특성을 분석하고, 이를 기반으로 정체성 정보와 포즈 정보를 각각 독립적인 참조 입력으로 활용하는 하이브리드 워크플로우를 설계하였다. 둘째, 동일한 인물 이미지와 목표 포즈 이미지를 기준으로 Identity, ReID, 하이브리드의 세 가지 방식에서 반복적으로 이미지를 생성하여 비교 가능한 실험 데이터를 구축하였다. 셋째, 생성 이미지와 기준 인물 간의 얼굴 특징 유사도를 기반으로 안면 정체성 유사

1) Ho, J., Jain, A. N., & Abbeel, P., 'Denoising Diffusion Probabilistic Models', NeurIPS, Advances in Neural Information Processing Systems, Vol.33, 2020. 12., p.6840.

2) Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B., 'High-Resolution Image Synthesis with Latent Diffusion Models', IEEE/CVF, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022. 06., p.10684.

3) Krea, Introducing Krea 2, (2026.08.19.)
www.Krea.ai/blog/Krea-2-image-model

4) GitHub, 'ComfyUI: A modular node-graph AI

creation engine', (2026.08.19.)
github.com/comfy-org/ComfyUI

도(Facial Identity Similarity, FIS)를 측정하고, 생성 이미지에서 추정된 신체 관절 좌표와 목표 포즈의 관절 좌표를 비교하여 포즈 정확도(Pose Accuracy, PA)를 산출하였다. 또한 동일 조건에서 반복 생성된 이미지 간의 얼굴 유사도를 비교하여 이미지 간 정체성 일관성(Inter-Image Identity Consistency, IIC)을 측정한다. 넷째, 포즈 변화 수준에 따른 FIS와 PA의 변화를 비교·분석함으로써 각 방식이 안면 정체성과 포즈 제어 사이에서 어떠한 특성과 상충관계를 나타내는지 분석하고자 한다. 이를 위해 본 연구는 다음과 같은 연구문제를 설정하였다.

연구문제 1. Identity와 ReID는 안면 정체성 유지와 포즈 제어에서 어떠한 차이를 보이는가?

연구문제 2. 정체성 참조와 포즈 참조를 기능적으로 구분한 하이브리드 워크플로우는 Identity 및 ReID와 비교하여 어떠한 성능 특성을 보이는가?

연구문제 3. 포즈 변화 수준이 높아질수록 Identity, ReID, 하이브리드 워크플로우의 FIS와 PA는 어떻게 변화하는가?

2. 이론적 배경

2-1. 생성형 AI 이미지의 인물 정체성과 포즈 제어

생성형 AI를 활용한 인물 이미지 제작에서는 확산 모델의 확률적 생성 특성으로 인해 동일한 입력 조건에서도 인물의 얼굴형과 이목구비 등 세부 특징이 달라질 수 있으며, 이에 따라 새로운 장면이나 자세에서도 동일한 인물의 정체성을 안정적으로 유지하는 것이 중요한 품질 요인으로 작용한다. 특히 캐릭터 시트, 광고 이미지, 스토리보드처럼 동일 인물이 반복적으로 등장하는 콘텐츠에서는 개별 이미지의 완성도뿐만 아니라 인물의 정체성과 신체 자세를 얼마나 안정적으로 제어할 수 있는지가 결과물의 활용 가능성을 결정한다.

본 연구에서 인물 정체성은 생성된 대상이 참조 인물과 동일한 사람으로 인식될 수 있도록 안면 특징이 유지되는 정도를 의미한다. 이러한 안면 정체성의 유지 정도를 정량적으로 분석하기 위해 본 연구에서는 ArcFace를 활용하였다. ArcFace는 정규화된 특징 공간에 각도 마진을 부여함으로써 동일 인물의 특징 표현은 조밀하게, 서로 다른 인물의 특징 표현은 분리되도록 학습된 대표적인 얼굴 인식 모델이다.⁵⁾ 이를 바

탕으로 기존 얼굴과 생성 얼굴에서 추출된 특징 임베딩 간의 코사인 유사도를 통해 참조 인물의 안면 특징이 생성 결과에서 어느 정도 유지되는지를 정량적으로 비교할 수 있다. 본 연구에서는 이를 안면 정체성 유사도(Facial Identity Similarity, FIS)로 정의하고, 생체적 동일성을 판정하는 절대적 기준이 아니라 동일한 전처리와 분석 조건에서 생성 방식 간 안면 정체성 유지 정도를 비교하기 위한 상대적 지표로 활용하였다.

포즈 제어는 생성된 인물의 몸통 방향과 사지 배치가 목표로 제시된 자세를 얼마나 정확하게 재현하는지를 의미한다. 2차원 자세 추정은 어깨, 팔꿈치, 손목, 엉덩이, 무릎 및 발목 등의 핵심점(keypoint)을 검출하여 인체 자세를 수치화하는 방식이다. OpenPose는 Part Affinity Fields를 이용하여 신체 부위 간의 연결 관계를 추정하는 대표적인 다중 인물 2차원 자세 추정 방법이다.⁶⁾ 본 연구에서는 참조 포즈와 생성 결과의 대응 핵심점 거리를 인물 크기에 따라 정규화하고, 일정 허용 범위 안에 위치한 핵심점의 비율을 측정하는 PCK(Percentage of Correct Keypoints)를 적용하여 포즈 정확도(Pose Accuracy, PA)를 산출한다. PCK는 핵심점 간 거리를 신체 크기 등의 기준으로 정규화한 뒤 설정된 임계값 이내에 위치한 핵심점의 비율을 산출하는 평가 방식이다.⁷⁾ MMPose에서는 PCK 계산 시 정규화된 거리의 임계값으로 0.05, 0.1, 0.2 등이 일반적으로 사용될 수 있음을 제시하고 있으며,⁸⁾ PCK@0.2는 기존 인체 자세 추정 연구에서도 활용되어 왔다.⁹⁾ 이때 얼굴 랜드마크는 안면 정체성을 평가

‘ArcFace: Additive Angular Margin Loss for Deep Face Recognition’, IEEE/CVF, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019. 06., p.4692.

- 6) Cao, Z., Simon, T., Wei, S.-E., & Sheikh, Y., ‘Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields’, IEEE, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017. 07., p.7294.
- 7) Andriluka, M., Pishchulin, L., Gehler, P., & Schiele, B., ‘2D Human Pose Estimation: New Benchmark and State of the Art Analysis’, IEEE, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014. 06., p.3690.
- 8) MMPose Documentation, [mmpose.evaluation.functional.keypoint_eval](https://mmpose.readthedocs.io/en/latest/_modules/mmpose/evaluation/functional/keypoint_eval.html), (2026.08.19.)
- 9) mmpose.readthedocs.io/en/latest/_modules/mmpose/evaluation/functional/keypoint_eval.html

5) Deng, J., Guo, J., Xue, N., & Zafeiriou, S.,

하는 FIS와 영역이 중복되고 작은 얼굴이나 가림에 민감하므로, PA에서는 몸통과 사지의 핵심점을 중심으로 평가한다.

하나의 생성 결과가 참조 인물과 얼마나 유사한지와, 동일한 생성 조건에서 반복된 결과 간에 인물의 정체성이 얼마나 일관되게 유지되는지는 서로 구분할 필요가 있다. 이에 본 연구에서는 동일 인물-생성 방식-포즈 조건에서 세 차례 생성된 결과로부터 세 개의 이미지 쌍을 구성하고, 각 쌍의 ArcFace 코사인 유사도를 산출한 뒤 그 평균을 이미지 간 정체성 일관성(IIC)으로 정의하였다. FIS가 참조 인물에 대한 정체성 유지 정도를 나타낸다면, IIC는 반복 생성 과정에서 동일한 인물 정체성이 얼마나 안정적으로 유지되는지를 나타낸다. 이상의 논의를 바탕으로 본 연구의 핵심 평가 개념과 조작적 정의를 정리하면 [표 1]과 같다.

한편 생성 조건에 DWPose를 활용한 경우, 동일한 자세 추정기로 생성 결과를 평가하면 생성과 평가 과정에 동일한 추정 특성이 반영되어 순환 평가의 문제가 발생할 수 있다. 이에 본 연구에서는 평가의 독립성을 확보하기 위해 MMPose 기반 자세 추정 결과를 평가에 활용하였다.¹⁰⁾

[표 1] 본 연구의 핵심 평가 개념과 조작적 정의

평가 개념	조작적 정의	주요 대상	해석
안면 정체성(FIS)	기준 생성 얼굴 임베딩 코사인 유사도	ArcFace 얼굴 특징	높을수록 참조정체성 유지
포즈 정확도(PA)	정규화 거리 기반 PCK@0.2	몸통-사지 핵심점	높을수록 목표 포즈 일치
이미지 간 정체성(IIC)	반복 결과의 쌍별 얼굴 유사도 평균	동일 조건 3회 시행 결과	높을수록 반복 생성 안정

2-2. Krea 2 Identity와 ReID

Krea 2는 창의적 탐색성과 제어 가능성을 강조한 이미지 생성 모델로, 텍스트와 이미지 참조를 활용하여 생성 결과의 방향을 조절할 수 있는 기반을 제공한다. 이러한 참조 기반 생성 방식은 동일 인물의 특징을 유

지하면서 새로운 장면이나 자세를 구현해야 하는 인물 이미지 제작에서 중요한 기술적 기반이 된다. 본 연구에서 다루는 Identity와 ReID는 Krea 2의 공식 표준 기능명이 아니라, 공개 커뮤니티에서 개발된 어댑터 및 ComfyUI 기반 워크플로우를 의미한다.

Krea 2 Identity Edit는 비공식 instruction-edit LoRA로, VAE 참조 토큰과 Qwen3-VL 이미지 조건화를 함께 활용하여 원본 또는 참조 인물의 특징을 유지하면서 이미지를 편집하거나 재배치하는 방식을 지향한다.¹¹⁾ 이 방식은 참조 이미지의 시각적 특징과 구조를 적극적으로 활용한다는 점에서 안면 정체성 보존에 유리할 가능성이 있으나, 원본과 크게 다른 자세를 구현할 때는 기존 구조에 대한 의존성이 제약으로 작용할 수 있다. 따라서 본 연구에서는 Identity를 비교 기준으로 설정하고, 안면 정체성 유지와 포즈 변화 수준에 따른 포즈 정확도의 특성을 실험적으로 분석하였다.

반면 Krea 2 ReID Reference는 한 장의 RGB 이미지를 인물의 정체성 참조로 활용하고, 출력 프롬프트를 통해 새로운 포즈와 장면을 생성하는 커뮤니티 기반 LoRA이다.¹²⁾ 참조 이미지를 최종 생성 구조와 상대적으로 분리된 정체성 조건으로 활용하기 때문에 새로운 자세를 구현하는 데 상대적으로 높은 자유도를 제공할 것으로 예상된다. 다만 안면 정체성의 보존 정도는 참조 이미지의 크롭, 프롬프트, 조건 강도 및 시드 등에 따라 달라질 수 있으므로, ReID의 정체성 유지 성능을 사전에 단정하기보다 안면 정체성 유사도(FIS)를 통해 정량적으로 분석할 필요가 있다. 또한 ReID는 Identity 생성 이후에 적용되는 후처리 모듈이 아니라, 독립적인 정체성 참조 기반 생성 경로로 구분하여 이해할 필요가 있다. 개발자가 제시한 기능적 특성에 따르면 Identity는 참조 이미지의 특징과 구조 보존에, ReID는 정체성 참조를 기반으로 한 새로운 포즈 및 장면 생성에 각각 강점을 갖는다. 이러한 기능적 차이를 정리하면 [표 2]와 같다.

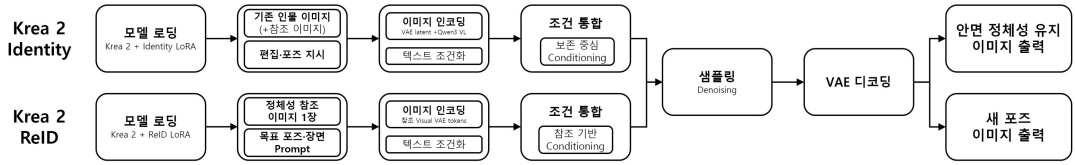
이처럼 Identity와 ReID는 어느 한 방식이 절대적으로 우수한 관계라기보다, 안면 정체성 유지와 포즈 제어라는 두 목표 사이에서 서로 다른 특성을 갖는 생성 방식으로 볼 수 있다. Identity는 참조 이미지의 특징

9) Zhang, F., Zhu, X., & Ye, M., 'Fast Human Pose Estimation', IEEE/CVF, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019. 06., p.3520.

10) GitHub, OpenMMLab Pose Estimation Toolbox and Benchmark, (2026.08.19.)
github.com/open-mmlab/mmpose

11) Hugging Face, Krea 2 Identity Edit v1.2, (2026.08.19.)
huggingface.co/Ty6666667/Krea2-identity-edit

12) Hugging Face, 'Krea 2 ReID Reference', (2026.08.19.)
huggingface.co/yijunwang2/Krea-2-reid



[그림 1] Krea 2 기반 Identity와 ReID의 생성 프로세스 비교

과 구조를 적극적으로 유지하는 반면, ReID는 정체성 정보를 독립적인 참조 조건으로 활용함으로써 새로운 포즈 생성에 상대적으로 높은 자유도를 제공한다.

이에 본 연구에서는 ReID 생성 결과에 Identity를 적용하는 직렬적 후처리 방식이 아니라, 포즈 정보와 정체성 정보의 역할을 기능적으로 분리한 하이브리드 워크플로우를 제안한다. 포즈 참조 이미지는 DWPose와 OpenPose Control LoRA를 통해 목표 자세가 반영된 포즈 기준 이미지로 변환되며, 이는 최종 Identity 기반 2-reference 생성 단계에서 'HOW' 정보로 활용된다. 정체성 참조 이미지는 동일한 단계에서 생성 인물의 안면 특징을 규정하는 'WHO' 정보로 활용되며, 두 정보는 서로 다른 참조 조건으로 결합되어 최종 이미지를 생성한다. 이를 통해 안면 정체성 유지와 포즈 변화 간의 상충관계가 하이브리드 방식에서 어떠한 양상으로 나타나는지를 Identity 및 ReID 기준군과의 비교를 통해 분석하고자 한다.

[표 2] Krea 2 Identity와 ReID의 기능적 비교

구분	Identity	ReID
주요 목적	참조 인물의 보존적 편집 및 재배치	정체성 참조 기반 새 포즈 생성
입력 방식	단일 편집 또는 장면1 + 인물1	정체성 참조 이미지 + 프롬프트
구조적 특성	원본 구조 보존 경향이 높음	출력 포즈 자유도가 높음
주요 강점	안면 정체성 보존	새 포즈로의 재생성
분석 과제	큰 포즈 변화 시 PA 저하 여부	조건 변화 시 FIS 저하 여부
본 연구에서의 역할	Identity 기준군 및 하이브리드의 2-reference 기반	독립 ReID 기준군

3. ComfyUI 기반 하이브리드 워크플로우 설계

3-1. Identity 및 ReID 워크플로우 분석

본 연구에서는 Krea 2 기반의 Identity와 ReID가 인물의 안면 정체성 유지와 포즈 제어에 어떠한 차이

를 보이는지 비교하기 위해 ComfyUI Desktop 환경에서 각 생성 경로의 기능적 구조를 분석하였다.

ComfyUI는 모델 로딩, 이미지 입력 및 인코딩, 텍스트 조건화, 샘플링, VAE 디코딩 등의 과정을 노드 그래프 형태로 구성할 수 있어, 각 참조 정보가 생성 과정에서 어떠한 역할을 수행하는지를 명시적으로 구분할 수 있다. 이에 본 연구에서는 전체 노드 구성을 그대로 제시하기보다, 각 워크플로우를 입력-조건화-생성-출력의 기능 단위로 단순화하여 비교하였다.

Identity 방식은 기존 이미지 또는 참조 인물의 시각적 특징을 유지하면서 새로운 결과를 생성하는 보존 중심의 조건화 경로를 갖는다.¹³⁾ 이 과정에서 참조 이미지의 안면 특징뿐만 아니라 기존 자세와 화면 구조가 생성 결과에 함께 영향을 미칠 가능성이 존재한다. 반면 ReID 방식은 단일 RGB 이미지를 인물의 정체성 참조로 활용하고, 프롬프트를 기반으로 새로운 자세와 장면을 구성하는 독립적인 정체성 참조 생성 경로를 갖는다.¹⁴⁾ 따라서 두 방식은 동일한 참조 인물을 활용 하더라도 정체성 정보와 포즈 정보가 생성 과정에 반영되는 방식에서 서로 다른 특성을 갖는다.

[그림 1]은 Identity와 ReID의 생성 과정을 기능 단위로 단순화한 구조이다. Identity는 기존 인물 이미지의 정보를 보존하는 조건화를 중심으로 생성이 이루어지는 반면, ReID는 정체성 참조 이미지를 인물의 기준 정보로 활용하면서 프롬프트와 모델의 생성 능력을 통해 새로운 자세를 구성한다. 이러한 구조적 차이는 Identity에서는 안면 정체성 보존이, ReID에서는 포즈 변화의 자유도가 상대적으로 높게 나타날 가능성을 제시한다. 다만 이는 각 방식의 구조적 특성에 따른 예상이며, 실제 성능은 이후 실험을 통해 정량적으로 비교·분석한다.

[표 3]은 두 방식의 성능을 공정하게 비교하기 위한 워크플로우의 분석 기준이다. 공정한 비교를 위해 출력

13) Hugging Face, 'Krea 2 Identity Edit v1.2', Op. cit., (2026.08.19.).

14) Hugging Face, 'Krea 2 ReID Reference', Op. cit., (2026.08.19.).

해상도, 샘플러, 스텝 수, 기본 모델 체크포인트, 프롬프트의 의미적 범위와 시드(seed) 집합을 동일하게 설정하였다. 또한 배경은 포즈 변화 수준을 결정하는 요인에서 제외하고 모든 생성 방식에서 동일하게 유지하였다. Identity의 보존 강도와 RelD의 LoRA 강도는 사전 파일럿 실험을 통해 결정한 뒤 전체 실험에 동일하게 적용하였으며, 얼굴 검출 또는 포즈 추정에 실패한 결과 역시 제외하지 않고 별도의 실패 사례로 기록하였다. 이를 통해 생성 방식의 구조적 차이에 따른 안면 정체성 유지와 포즈 제어 성능의 차이를 비교할 수 있도록 실험 조건을 통제하였다.

[표 3] 워크플로우 분석 기준

분석 요소	Identity	RelD	실험 통제
참조 역할	기준 인물 원본 보존	정체성 앵커	동일 인물 참조 사용
포즈 정보	원본 구조 영향 가능	프롬프트와 생성 prior에 의존	목표 포즈를 조건별 고정
프롬프트	편집 지시 중심	출력 장면 기술 중심	의미적 목표 및 기술 범위 통일
출력 특성	보존 강도의 영향을 받음	참조 크롭 및 LoRA 강도에 영향을 받음	해상도·seed·step 고정

3-2. 정체성 및 포즈 정보 분리 기반 하이브리드 워크플로우 설계

본 연구에서 제안하는 하이브리드 워크플로우의 핵심은 하나의 참조 이미지가 인물의 정체성과 포즈 정보를 동시에 지배하도록 하지 않고, 두 정보의 역할을 기능적으로 분리하는 데 있다. 이를 위해 포즈 참조(pose reference)는 생성될 인물의 목표 자세, 몸통 방향 및 사지 배치를 결정하는 ‘HOW’ 정보로 설정하고, 정체성 참조(identity reference)는 기준 인물의 얼굴형과 이목구비 등 안면 정체성을 규정하는 ‘WHO’ 정보로 구분하였다. 포즈 참조 이미지는 DWPose 기반의 포즈 생성 경로를 통해 목표 자세가 반영된 포즈 기준 이미지로 변환되며, 이 결과는 최종 Identity 기반 2-reference 생성 단계에서 ‘HOW’ 정보로 활용된다. 정체성 참조 이미지는 동일한 최종 생성 단계에서 기준 인물의 안면 특징을 규정하는 ‘WHO’ 정보로 입력되며, 두 정보는 서로 다른 참조 역할로 결합된다. 이는 모델 내부의 표현 공간을 수학적으로 분해하는 방식이 아니라, 워크플로우의 입력과 조건화 단계에서

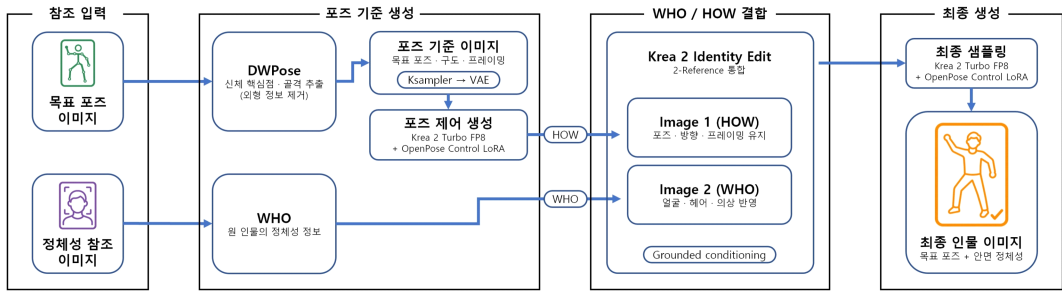
각 참조 이미지의 기능을 의도적으로 구분하는 접근이다. 이를 통해 포즈 참조가 정체성 정보에 미치는 영향과 정체성 참조가 목표 자세에 미치는 영향을 최소화하고, 각 정보원이 담당하는 기능을 명확히 하고자 하였다.

[그림 2]와 같이 제안된 구조에서는 먼저 DWPose를 이용하여 포즈 참조 이미지의 신체 핵심점과 골격 구조를 추출한다. 추출된 포즈 맵은 Krea 2 OpenPose Control LoRA의 제어 입력으로 사용되며, 이를 통해 목표 자세와 신체 배치가 반영된 포즈 기준 이미지를 생성한다. 이후 포즈 기준 이미지는 ‘HOW’ 정보로, 정체성 참조 이미지는 ‘WHO’ 정보로 각각 최종 Identity 기반 2-reference 생성 단계에 입력되어 하나의 결과 이미지로 통합된다. 이 과정에서 포즈 기준 이미지가 중간 단계에서 생성되지만, 이는 RelD를 통해 완성 이미지를 생성한 후 Identity를 적용하여 안면을 다시 보정하는 직렬적 후처리 방식과는 구분된다. 포즈 기준 이미지는 최종 결과물을 수정하기 위한 대상이 아니라, 목표 자세를 최종 생성 단계에 전달하기 위한 중간 정보원으로 활용된다.



[그림 2] 정체성-포즈 정보의 기능적 분리 기반의 하이브리드 워크플로우 개념도

본 연구가 제안하는 하이브리드 방식의 목적은 안면 정체성 유사도(FIS)와 포즈 정확도(PA) 가운데 하나의 성능을 극대화하는 것이 아니라, Identity 방식에서 기대되는 안면 정체성 유지와 RelD 방식에서 기대되는 포즈 변화의 유연성을 하나의 워크플로우에서 기능적으로 분리·결합함으로써 두 목표 사이에서 발생하는 상충 관계를 완화하는 데 있다. 따라서 하이브리드 방식의 성능 특성은 모든 평가지표에서 가장 높은 값을 나타내지보다, Identity의 정체성 유지 특성과 RelD의 포즈 재현 특성을 어느 정도 함께 나타내는지를 중심으로 비교·분석하였다.



[그림 3] 정체성-포즈 정보가 결합된 하이브리드 워크플로우 개요도

이를 위한 하이브리드 워크플로우의 설계 원칙은 다음과 같이 설정하였다. 첫째, 포즈 참조는 HOW, 정체성 참조는 WHO로 역할을 고정하여 두 정보원의 기능을 명확히 구분한다. 둘째, 포즈 참조 이미지는 원본 RGB 정보를 직접 정체성 조건으로 사용하지 않고, DWPose를 통해 핵심점 기반의 자세 정보로 변환하여 포즈 생성 경로에 활용함으로써 참조 인물의 외형 및 인면 정체성 정보가 직접적으로 유입되는 것을 최소화한다. 셋째, Identity, ReID, 하이브리드의 비교 과정에서는 동일한 인물, 생성 조건 및 시드(seed) 집합을 적용하여 방식 간 비교 가능성을 확보한다. 넷째, ComfyUI 워크플로우 JSON, 모델 및 노드 버전, seed와 프롬프트를 함께 기록하여 실험 과정의 재현성을 확보한다. 설계원칙에 따라 ComfyUI에서 설계한 하이브리드 워크플로우를 표현하면 [그림 3]과 같다.

3-3. 실험 설계 및 평가방법

모든 이미지 생성 실험은 동일한 하드웨어 및 소프트웨어 환경에서 수행하였다. [표 4]는 실험의 재현성을 확보하기 위해 ComfyUI 환경, 모델 및 LoRA 체크포인트, 출력 해상도, 샘플링 스텝, sampler, seed 등의 주요 생성 조건을 기록한 내용이다. 이러한 통제 조건을 바탕으로 Identity, ReID 및 하이브리드 워크플로우의 인면 정체성 유지와 포즈 제어 성능을 비교하였다. 실험은 5개의 합성 인물 정체성, 3단계의 포즈 변화 수준, 3개의 생성 방식과 조건별 3회 반복으로 구성하였으며, 총 135장(5×3×3)의 이미지를 생성하였다. 실제 인물의 초상권 및 개인정보 문제를 최소화하기 위해 생성형 AI로 제작한 가상의 인물을 실험 대상으로 사용하였으며, 기준 얼굴 이미지는 경면의 단일 인물로 구성하고 충분한 해상도와 얼굴 가림이 없는 조건으로 통일하였다.

[표 4] 실험 설정 및 재현 환경

항목	실험 설정
Hardware / OS	NVIDIA RTX 4080 / Windows 11
Environment	ComfyUI Desktop v.0.33.2
Base Model	Krea 2 Turbo FP8
Identity LoRA	Krea 2 Identity Edit v1.2
Pose Condition	DWPose + Krea 2 OpenPose Control LoRA
Resolution	1:1, 1 MP (1024*1024px)
Sampling	Euler, 10 steps, CFG 1.0, Simple
Repetition	1차:13579 / 2차:24680 / 3차:12345
ReID LoRA	Krea 2 ReID Reference
ReID Strength	1.0

생성 방식에 따른 차이를 명확하게 확인하기 위해 앞서 제시한 [표 3] 워크플로우 분석 기준에 따라 변인을 통제하였다. 또한 포즈 변화 수준은 기준 자세와 생성 목표 간의 몸통 방향 및 사지 배치 차이를 중심으로 설정하였으며, 배경·조명·의상 등은 포즈 변화 수준의 결정 요인에서 제외하고 동일하게 유지하였다. 다만 Identity, ReID 및 하이브리드는 참조 정보의 입력 방식과 조건화 구조, 생성 단계 수가 서로 다르므로 각 방식의 고유한 생성 구조는 그대로 유지하였다. 따라서 본 실험은 입력 구조와 생성 단계를 동일화한 모델 단위의 비교가 아니라, 기준 인물, 목표 포즈, 출력 해상도 및 주요 샘플링 조건을 통제한 워크플로우 수준의 탐색적 비교로 수행하였다. 특히 하이브리드 방식은 DWPose와 OpenPose Control LoRA를 활용한 포즈 기준 이미지의 선형 생성과 Identity 기반 2-reference 결합 단계를 포함한다. 이에 따라 관찰된 성능 차이는 기능적 참조 분리, 추가 포즈 조건화 및 2단계 생성구조가 결합된 결과로 보아야 하며, 기능적 참조 분리만의 독립적 효과로 해석할 수 없다. 동일한 목표 포즈는 공통의 목표 자세를 기준으로 설정하되, 각 방식의 고유한 입력 구조에 따라 이미지 참조 또는 이에 대응하는 프롬프트 조건으로 적용하였다.

포즈 변화 수준은 기준 이미지와 목표 자세 간 구조적 차이를 기준으로 Low, Medium, High의 세 단계로 구분한다. Low는 정면 방향을 유지하면서 팔이나 다리의 위치가 소폭 변화하는 수준이며, Medium은 정면에서 3/4 방향으로의 회전, 사지 배치의 변화 및 가벼운 걷기 자세 등을 포함한다. High는 측면 방향으로의 큰 회전, 앞가달리가-몸통 비틀기와 같이 신체 구조가 크게 변화하는 자세를 의미한다. 이때 장면이나 배경의 변화는 포즈 변화 수준의 분류 기준에 포함하지 않는다.

안면 정체성 유지 성능은 안면 정체성 유사도(Facial Identity Similarity, FIS)를 통해 측정하였다. 기준 얼굴과 생성 얼굴에서 ArcFace 임베딩을 추출한 뒤 두 임베딩 간 코사인 유사도를 산출하여 생성 이미지가 참조 인물의 안면 특징을 어느 정도 유지하는지를 평가하였다.¹⁵⁾ 분석 결과는 얼굴이 정상적으로 검출된 표본을 대상으로 평균과 표준편차를 산출하고, 유효 표본 수(n)를 함께 제시하였다.

포즈 재현 성능은 포즈 정확도(Pose Accuracy, PA)를 통해 측정하였다. 참조 포즈와 생성 결과에서 어깨, 팔꿈치, 손목, 엉덩이, 무릎 및 발목의 대응 핵심점을 추출한 뒤, 인물의 bounding box 너비와 높이를 기준으로 x축과 y축의 핵심점 좌표 차이를 각각 정규화하였다. 본 연구에서는 정규화된 거리가 0.2 이하인 핵심점을 올바르게 재현된 것으로 판단하는 PCK@0.2를 적용하고,¹⁶⁾ 전체 유효 핵심점 가운데 해당 기준을 만족한 핵심점의 비율을 PA로 산출하였다. 얼굴 핵심점은 FIS와 평가 영역이 중복되므로 분석 대상에서 제외하였다. 생성 과정의 포즈 조건에는 DWPose를 활용하였으며¹⁷⁾ 동일한 추정기를 평가에 다시 적용할 경우 평가 편향이 발생할 가능성이 있으므로, 평가에는 생성 경로와 독립된 MMPose를 활용하였다.¹⁸⁾

15) Deng, J., Guo, J., Xue, N., & Zafeiriou, S., Op. cit. 2019, p.4692.

16) MMPose Documentation, 'mmpose.evaluation.functional.keypoint_eval', Op. cit., (2026.08.19.).

17) Yang, Z., Zeng, A., Yuan, C., & Li, Y., 'Effective Whole-Body Pose Estimation with Two-Stages Distillation', IEEE/CVF, Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, 2023. 10., p.4210.

18) OpenMMLab GitHub, 'MMPose: OpenMMLab Pose Estimation Toolbox and Benchmark', Op. cit., (2026.08.19.).

반복 생성에 따른 정체성의 안정성은 이미지 간 정체성 일관성(Inter-Image Identity Consistency, IIC)을 통해 평가한다. 동일한 인물, 생성 방식 및 포즈 수준에서 세 차례 생성된 이미지의 모든 쌍에 대해 ArcFace 코사인 유사도를 산출하고 그 평균을 구한다. 이를 통해 FIS가 참조 인물과 생성 결과 사이의 정체성 유지 정도를 측정한다면, IIC는 동일한 조건에서 반복 생성된 결과들이 얼마나 일관된 인물 정체성을 유지하는지를 평가한다.

최종적으로 FIS, PA 및 IIC를 중심으로 생성 방식과 포즈 변화 수준에 따른 결과의 차이를 기술통계적으로 비교·분석하였다. FIS와 PA는 각 생성 방식 및 포즈 수준별 평균과 표준편차를 산출하여 전반적인 성능 수준과 변동 양상을 비교하였으며, IIC는 동일한 인물 생성 방식·포즈 조건에서 반복 생성된 결과 간의 평균 유사도를 기준으로 생성 안정성을 분석하였다. 또한 Low, Medium, High의 포즈 변화 수준에 따른 FIS와 PA의 증감 양상을 비교하여 각 생성 방식에서 나타나는 안면 정체성 유지와 포즈 제어 특성을 탐색적으로 분석하였다. 반복 생성된 이미지는 동일 인물과 포즈 조건에 내재된 관측치이므로 서로 완전히 독립된 표본으로 간주하지 않았다. 이에 본 연구는 제한된 수의 합성 인물과 생성 조건을 대상으로 한 탐색적 워크플로우 비교연구로서, 관찰된 차이를 모집단 수준의 통계적 차이로 일반화하기보다 생성 방식 간 상대적 성능과 변화 양상을 파악하는 데 분석의 목적을 두었다.

4. 실험 결과 및 분석

4-1. 생성 방식에 따른 안면 정체성 및 포즈 제어 비교

Identity, RelD, 하이브리드 방식의 안면 정체성 유사도(FIS), 포즈 정확도(PA), 생성 이미지 간 정체성 일관성(IIC)을 비교한 결과는 [표 5]와 같다. 각 방식은 5개의 합성 인물과 3단계의 포즈 변화 수준을 대상으로 조건별 3회 반복 생성하여 평가하였다.

분석 결과, Identity 방식의 FIS는 0.82, IIC는 0.79로 나타나 기준 인물의 안면 특징과 반복 생성 간 정체성을 비교적 안정적으로 유지하는 경향을 보였다. 반면 PA는 0.46으로 세 방식 중 가장 낮게 나타났다. 실제 생성 결과에서도 얼굴형과 이목구비는 기준 이미지와 유사하게 유지되었으나, 포즈 참조에 제시된 몸통 방향이나 사지 배치가 충분히 반영되지 않고 기준 이미지의 자세에 가까운 결과가 생성되는 사례가 확인되

었다. 이는 Identity 방식이 안면 정체성 유지에는 상대적으로 강점을 보이는 반면, 큰 포즈 변화를 구현하는 과정에서는 구조적 제약이 발생할 수 있음을 보여준다.

[표 5] 생성 방식별 정량 평가 결과(평균±표준편차)

생성방식	FIS	PA	IIC	유효 n (FIS,PA)	유효조건 n (IIC)
Identity	0.82 ±0.06	0.46 ±0.15	0.79 ±0.07	45	15
RelD	0.58 ±0.09	0.78 ±0.11	0.61 ±0.10	45	15
하이브리드	0.87 ±0.04	0.91 ±0.06	0.85 ±0.05	45	15

RelD 방식은 PA가 0.78로 나타나 Identity보다 높은 포즈 제어 성능을 보였다. 특히 기준 자세와 다른 신체 자세를 구현하는 데 상대적으로 높은 자유도를 보였다. 그러나 FIS는 0.58, IIC는 0.61로 Identity보다 낮게 나타났다. 생성 이미지가 참조 인물과 전반적으로 유사한 인상을 형성하더라도 얼굴 윤곽이나 이목구비의 비율이 달라지는 사례가 나타났으며, 반복 생성 간에도 동일한 안면 특징이 안정적으로 유지되지 않는 경향이 확인되었다. 따라서 RelD 방식은 새로운 포즈를 구현하는 데 상대적으로 높은 수준을 보였으나, 참조 인물의 안면 정체성 유지에서는 Identity와 하이브리드보다 상대적으로 낮은 수준을 보였다.

하이브리드 방식은 FIS 0.87, PA 0.91, IIC 0.85로 본 실험에서 관찰된 세 방식의 평균값 가운데 모두 가장 높게 나타났다. Identity와 비교하면 FIS는 0.05, PA는 0.45, IIC는 0.06 높았으며, RelD와 비교하면 FIS는 0.29, PA는 0.13, IIC는 0.24 높게 나타났다. 특히 PA가 0.91로 나타난 동시에 Identity보다 높은 FIS와 IIC가 관찰되어, 본 실험 조건에서 안면 정체성 유지와 포즈 제어가 비교적 균형적으로 나타나는 성능 특성을 보였다. 다만 하이브리드 방식은 기능적 참조 분리뿐 아니라 별도의 포즈 조건화와 2단계 생성구조를 포함하므로, 이러한 차이를 기능적 참조 분리만의 독립적 효과로 해석할 수는 없다.

또한 하이브리드의 표준편차는 FIS 0.04, PA 0.06, IIC 0.05로 세 방식 가운데 상대적으로 낮게 나타났다. 이는 본 연구의 실험 조건에서 하이브리드 방식의 결과 변동성이 다른 방식에 비해 상대적으로 작게 나타났음을 보여준다. 따라서 세 지표에서 나타난 상대적으로 높은 평균값과 낮은 변동성은 기능적 참조 분리, 추가 포즈 조건화 및 2단계 생성구조가 결합된 하이브리드

드 워크플로우 전체의 성능 특성으로 해석하였다. 다만 [표 5]는 생성 방식별 전체 평균을 제시한 결과이므로, 포즈 변화 수준이 증가함에 따라 각 방식의 성능이 어떠한 양상으로 변화하는지는 직접적으로 확인하기 어렵다. 이에 본 연구에서는 전체 평균에서 드러나지 않는 성능 변화 양상을 구체적으로 검토하기 위해 포즈 변화 수준을 Low, Medium, High로 구분하고, 각 수준에 따른 FIS와 PA의 변화를 추가적으로 분석하였다.

4-2. 포즈 변화 수준에 따른 결과

[표 6]은 앞서 설정한 Low, Medium, High의 포즈 변화 수준에 따른 생성 방식별 FIS와 PA의 결과이다. Low는 기준 자세와 유사한 몸통 방향에서 사지 위치가 일부 변화하는 조건이며, Medium은 3/4 방향의 몸통 회전과 사지 배치 변화 및 가벼운 걷기 자세를 포함한다. High는 측면 회전, 앉기, 달리기 및 몸통 비틀기와 같이 기준 자세와의 구조적 차이가 크게 나타나는 조건으로 구성하였다.

[표 6] 변화 수준별 핵심 지표(평균±표준편차)

생성 방식	FIS			PA		
	Low	Med.	High	Low	Med.	High
Identity	0.87 ±0.03	0.83 ±0.05	0.76 ±0.07	0.78 ±0.07	0.42 ±0.11	0.18 ±0.09
RelD	0.64 ±0.07	0.58 ±0.08	0.52 ±0.09	0.85 ±0.06	0.79 ±0.08	0.70 ±0.12
하이브리드	0.90 ±0.03	0.88 ±0.04	0.84 ±0.05	0.95 ±0.03	0.92 ±0.04	0.86 ±0.06

Identity 방식의 FIS는 Low 0.87, Medium 0.83, High 0.76으로 나타나 포즈 변화 수준이 증가하더라도 비교적 높은 안면 정체성 유지 성능을 보였다. 반면 PA는 Low 0.78에서 Medium 0.42, High 0.18로 급격하게 감소하였다. Low 조건에서는 작은 사지 변화와 유사한 몸통 방향이 비교적 안정적으로 구현되었으나, Medium 이후부터는 기준 이미지의 구조적 영향이 강하게 나타났다. 특히 High 조건에서는 목표 자세가 충분히 반영되지 않거나 기준 자세에 가까운 형태로 생성되는 사례가 증가하였다. Low와 High 간 PA 감소폭은 0.60으로 세 방식 중 가장 크게 나타났으며, 이는 포즈 변화 수준이 증가함에 따라 Identity 방식의 포즈 정확도가 크게 감소하는 경향이 나타났음을 보여준다.

RelD 방식의 PA는 Low 0.85, Medium 0.79, High 0.70으로 나타났다. 포즈 변화 수준이 증가함에

따라 점수가 감소하였으나, High 조건에서도 상대적으로 높은 포즈 정확도를 유지하였다. 이는 ReID 방식에서 기준 자세와의 차이가 큰 조건에서도 새로운 포즈를 비교적 안정적으로 재현하는 양상이 나타났음을 보여준다. 반면 FIS는 Low 0.64, Medium 0.58, High 0.52로 모든 조건에서 Identity와 하이브리드보다 낮게 나타났다. 특히 포즈 변화 수준이 증가할수록 얼굴 윤곽과 이목구비의 변화가 두드러지는 사례가 증가하였다. 따라서 ReID는 포즈 변화에 대한 대응력은 상대적으로 높게 나타났으나, 참조 인물의 안면 정체성 유지 정도에서는 Identity와 하이브리드에 비해 상대적으로 낮은 FIS를 보였다. 하이브리드 방식의 FIS는 Low 0.90, Medium 0.88, High 0.84로 나타났으며, PA는 각각 0.95, 0.92, 0.86으로 측정되었다. Low에서 High로 변화했을 때 FIS의 감소폭은 0.06, PA의 감소폭은 0.09로 세 방식 중 가장 작았다. 특히 High 조건에서도 기준 인물의 얼굴형과 이목구비가 비교적 안정적으로 유지되었으며, 포즈 참조에서 제시한 몸통 방향과 사지 배치 역시 높은 수준으로 구현되었다.

방식 간 차이는 포즈 변화 수준이 증가할수록 더욱 뚜렷하게 나타났다. Medium 조건에서 하이브리드의 PA는 0.92로 Identity의 0.42보다 0.50 높았으며, FIS는 0.88로 ReID의 0.58보다 0.30 높았다. High 조건에서도 하이브리드의 PA는 Identity보다 0.68 높았고, FIS는 ReID보다 0.32 높게 나타났다. 이러한 결과는 포즈 변화 수준이 증가하는 조건에서 하이브리드 방식과 비교 방식 간 FIS 및 PA의 수치적 차이가 더욱 크게 나타나는 양상을 보여준다.

종합하면 Identity는 포즈 변화 수준이 증가하더라도 안면 정체성을 비교적 안정적으로 유지하였으나 포즈 정확도가 큰 폭으로 감소하였으며, ReID는 포즈 정확도를 상대적으로 안정적으로 유지한 반면 안면 정체성 유사도가 전반적으로 낮게 나타났다. 이에 비해 하이브리드는 포즈 변화 수준이 증가한 조건에서도 본 실험에서 관찰된 FIS와 PA 평균값이 두 비교 방식보다 상대적으로 높게 나타났다. 다만 이러한 양상은 정체성 참조와 포즈 참조의 기능적 분리뿐 아니라 DWPose와 OpenPose Control LoRA를 활용한 추가 포즈 조건화 및 2단계 생성구조가 함께 적용된 결과이다. 따라서 본 결과는 기능적 참조 분리만의 독립적 효과나 상충 관계의 일반적 완화로 해석하기보다, 제한된 실험 조건에서 확인된 하이브리드 워크플로우 전체의 탐색적 성능 특성으로 해석할 필요가 있다.

5. 결론

본 연구에서는 Krea 2 기반 Identity와 ReID 생성 방식의 안면 정체성 유지 및 포즈 제어 특성을 비교하고, 정체성 참조와 포즈 참조의 역할을 기능적으로 분리한 ComfyUI 기반 하이브리드 워크플로우를 제안하였다. 제안한 워크플로우의 성능 특성을 살펴보기 위해 동일한 인물과 포즈 조건에서 세 가지 생성 방식을 비교하고, 안면 정체성 유사도(FIS), 포즈 정확도(PA), 이미지 간 정체성 일관성(IIC)을 중심으로 결과를 정량적으로 비교·분석하였다.

분석 결과, Identity와 ReID는 안면 정체성 유지와 포즈 제어 측면에서 서로 다른 특성을 보였다. Identity 방식은 기준 인물의 얼굴형과 이목구비를 비교적 안정적으로 유지하였으나, 포즈 변화 수준이 높아질수록 기준 이미지의 구조적 영향이 강하게 나타나 목표 자세의 구현 정확도가 낮아지는 경향을 보였다. 반면 ReID 방식은 새로운 포즈를 생성하는 데 상대적으로 높은 자유도를 보였으며, 큰 포즈 변화에서도 목표 자세를 비교적 안정적으로 구현하였다. 그러나 얼굴형과 이목구비의 세부적인 관계가 변화하거나 반복 생성 간 얼굴 특징이 일정하게 유지되지 않는 사례가 확인되어, 안면 정체성 보존 측면에서는 상대적인 한계를 보였다.

본 연구에서 제안한 하이브리드 방식은 정체성 참조를 생성될 인물의 특징을 규정하는 'WHO' 정보로, 포즈 참조로부터 생성된 포즈 기준 이미지를 신체 자세를 전달하는 'HOW' 정보로 기능적으로 분리하고, 두 정보를 최종 Identity 기반 2-reference 생성 단계에서 서로 다른 참조 조건으로 통합하였다. 본 실험 조건에서 하이브리드 방식은 FIS, PA 및 IIC에서 비교적 높은 평균값과 낮은 변동성을 보였으며, 특히 포즈 변화 수준이 높아지는 조건에서도 FIS와 PA의 감소폭이 두 비교 방식보다 상대적으로 작게 나타났다. 다만 하이브리드 방식은 기능적 참조 분리뿐 아니라 DWPose와 OpenPose Control LoRA를 활용한 추가 포즈 조건화 및 2단계 생성구조를 포함한다. 따라서 관찰된 차이는 기능적 참조 분리만의 독립적 효과가 아니라, 복수의 구성요소가 결합된 하이브리드 워크플로우 전체의 탐색적 성능 특성으로 해석할 필요가 있다.

이러한 결과는 생성형 AI 기반 인물 이미지 제작에서 모델이나 LoRA의 선택뿐만 아니라, 참조 정보의 역할을 어떻게 구조화하고 분리할 것인지 역시 중요한 워크플로우 설계 요소가 될 수 있음을 시사한다. 특히 ComfyUI와 같은 노드 기반 환경에서는 정체성, 포즈, 장면 등 서로 다른 조건 정보의 기능을 명시적으로 구

분하여 생성 과정에 적용할 수 있다는 점에서, 동일 인물이 반복적으로 등장하는 캐릭터 시트, 광고 시안, 스토리보드 등의 제작 과정에 활용 가능성이 있다.

다만 본 연구는 5개의 합성 인물과 제한된 포즈 조건을 대상으로 수행되었으며, ArcFace와 자세 추정 기반의 자동 평가지표를 중심으로 분석하였다는 한계를 갖는다. 또한 제한된 표본과 생성 조건에서 방식 간 상대적 성능과 변화 양상을 살펴본 탐색적 비교연구이므로, 관찰된 수치 차이를 모집단 수준의 통계적 차이로 일반화하기 어렵다. 아울러 하이브리드 워크플로우는 기능적 참조 분리, 추가 포즈 조건화 및 2단계 생성구조를 함께 포함하므로, 각 구성요소의 영향을 개별적으로 분리하여 검증하지 못하였다. 따라서 본 연구의 결과는 제한된 조건에서 확인된 워크플로우 수준의 성능 특성으로 해석할 필요가 있으며, 자동 평가 결과가 사용자가 지각하는 인물의 동일성이나 포즈의 자연스러움과 완전히 일치한다고 단정하기도 어렵다. 더불어 지면상의 제약으로 동일한 정체성 참조와 목표 포즈에 따른 방식별 생성 결과를 시각적 비교표로 충분히 제시하지 못하여, 정량지표와 실제 시각적 차이의 관계를 직접 확인하는 데 한계가 있다. 후속 연구에서는 직접 포즈 조건화 대조군과 구성요소 제거 조건을 포함한 절제 실험(ablation study)을 통해 기능적 참조 분리와 추가 조건화의 영향을 구분하여 검증할 필요가 있다. 이와 함께 인물과 포즈 표본 및 생성 모델을 확대하고, 방식별 대표 결과와 실패 사례를 포함한 시각적 비교자료를 제시하는 한편, 실제 인물의 동의에 기반한 데이터와 인간 지각 평가 및 연구자 기반의 휴리스틱 평가를 병행함으로써 하이브리드 워크플로우의 성능 특성과 적용 가능성을 더욱 종합적으로 분석할 필요가 있다.

참고문헌

1. Andriluka, M., Pishchulin, L., Gehler, P., & Schiele, B., '2D Human Pose Estimation: New Benchmark and State of the Art Analysis', IEEE, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014.

2. Cao, Z., Simon, T., Wei, S.-E., & Sheikh, Y., 'Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields', IEEE, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.
3. Deng, J., Guo, J., Xue, N., & Zafeiriou, S., 'ArcFace: Additive Angular Margin Loss for Deep Face Recognition', IEEE/CVF, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019.
4. Ho, J., Jain, A. N., & Abbeel, P., 'Denoising Diffusion Probabilistic Models', NeurIPS, Advances in Neural Information Processing Systems, Vol.33, 2020.
5. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B., 'High-Resolution Image Synthesis with Latent Diffusion Models', IEEE/CVF, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022.
6. Yang, Z., Zeng, A., Yuan, C., & Li, Y., 'Effective Whole-Body Pose Estimation with Two-Stages Distillation', IEEE/CVF, Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, 2023.
7. Zhang, F., Zhu, X., & Ye, M., 'Fast Human Pose Estimation', IEEE/CVF, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019.
8. github.com
9. huggingface.com
10. mmpose.readthedocs.io
11. www.krea.ai