**TryOnDiffusion: A Tale of Two UNets 전처리 과정**

인물과 옷 이미지에서 파싱 맵과 포즈 키포인트를 획득하고 복잡한 옷 및 포즈 키포인트을 정확하게 추출하는 전처리 과정 방법을 논문 토대로 정리

1. **인체와 의복 이미지로부터 파싱 맵과 포즈 키포인트를 추출하여 다양한 부위와 위치 정보를 확보합니다.**

* Sp, Sg, Jp, Jg를 예측하여 인물과 옷 이미지에서 파싱 맵과 포즈 키포인트를 획득

|  |  |
| --- | --- |
| Sp  (separated person) | 인간 파싱 맵, 이미지 내의 인 형상의 다양한 부분을 나타내는 레이블 정보를 포함하는 맵 (머리, 상체, 하체, 팔, 다리 등의 부위를 식별하는 역할 수행) |
| Sg  (separated  Garment) | 옷 파싱 맵, 이미지 내의 의복의 다양한 부분을 나타내는 레이블 정보를 포함하는 맵으 옷의 종류와 부위를 식별하는 데 사용 |
| Jp  (joint  Person) | 2D 포즈 키포인트, 인물 이미지에서 특정 부위의 위치를 나타내는 포즈 키포인트 정보 |
| Jg  (joint  Germent) | 옷 2D 포즈 키포인트, 옷 이미지에서 특정 부위의 위치를 나타내는 포즈 키포인트 정보 |

* 1. **Sp,Sg,Jp,Jg 예측 방법**

1. Sp와 Sg의 경우 : 휴먼 파싱(human parsing)

|  |
| --- |
| * **휴먼 파싱의 정의?**   휴먼 파싱(Human Parsing)은 인간의 이미지나 영상에서 신체 부위와 의복 부위를 픽셀 단위로 분할하여 각 부위의 레이블 정보를 추출하는 작업을 의미   * 휴먼 파싱 방식 중 ‘Graphonomy: Universal Human Parsing via Graph Transfer Learning‘ 논문에 나오는 ‘Graphonomy’방식 이용 * ‘Graphonomy’ 작동원리 : 2가지 모듈로 구성돼어 각각 그래프 내 추론, 그래프 간 전이 역할을 수행한다.   (1)Intra-Graph Reasoning (그래프 내 추론):   * 이미지를 그래프로 변환 * 각 픽셀은 노드, 관계는 엣지로 표현 * 픽셀 간의 관계를 활용하여 자세한 분할 * 인접한 픽셀 정보로 정확한 분할 및 레이블 할당   (2)Inter-Graph Transfer (그래프 간 전이):   * 다른 데이터셋의 그래프 가져옴 * 이미지를 그래프로 표현하여 새 데이터셋에 적용 * 학습된 그래프 지식 전이 * 새 데이터셋에서도 효과적인 휴먼 파싱 수행 |

(2) Jp와 Jg의 경우 : 포즈 추정(pose estimation)

|  |
| --- |
| **- 포즈 추정의 정의?**  주어진 이미지나 영상에서 사람이나 객체의 키포인트(관절) 위치를 예측하는 기술  - 포즈 추정 방식 중 ‘Towards Accurate Multi-person Pose Estimation in the Wild’논문에 나오는 ‘multi-person pose estimation’ 방식사용  - multi-person pose estimation 의 작동 방식  (1) Faster RCNN 디텍터와 ResNet 활용:   * 사람 존재 가능성 있는 박스 위치 및 크기 예측 (Faster RCNN) * 각 박스에 포함될 가능성 있는 사람의 키포인트 추정 (ResNet)  1. 키포인트 추정 방식:  * 밀도 높은 키포인트 히트맵 및 오프셋 예측 (ResNet) * 키포인트 유형별로 예측 결과 수집 * 새로운 집계 절차 도입하여 국부화된 키포인트 예측 획득  1. 성능 향상 및 다양한 적용:  * 키포인트 기반 NMS 및 신뢰도 점수 계산 활용하여 정확한 결과 도출 * COCO keypoints 데이터셋에서 최첨단 결과 증명 |

1. **옷 분할 이미지(Ic)와 인물 분할 이미지(Ia)를 생성하여 의복 부분과 인체 형상을 따로 추출하기**
2. Ic (Isolated clothing, 옷 RGB 이미지): 옷 이미지에서 실제 의복 부분을 따로 분할한 이미지. 옷 파싱 맵(Sg)을 사용하여 의복 부분을 분리하여 나타낸 것으로, 옷을 개별적으로 조작할 때 사용
3. Ia (Isolated atomy, 인물 RGB 이미지): 인물 이미지에서 옷 부분을 제거하고, 개인의 형상만을 보존한 이미지. 인물 파싱 맵(Sp)을 사용하여 옷 부분을 없앤 이미지로, 옷을 입지 않은 인물 이미지로 시착 결과의 정확성을 높이기 위해 사용

**2-1. Ia 생성 원리**

* 인물 이미지에서 옷이 입혀져 있는 부분을 바운딩 박스로 마스킹 처리하여 가려준 후 다시 그 위에 머리, 손, 다리 같은 신체 부분을 복사해서 붙여 넣는다. -> 원래 옷의 정보는 완전히 사라짐!
* Ia의 필요성 : 어려운 인간 포즈나 느슨한 의복과 같이 복잡한 상황에서 의복 정보 노출을 줄이고 시착 결과의 정확성을 향상시키는 데 사용. 어려운 자세나 느슨한 옷을 입은 경우에도 옷 정보를 최대한 없앰으로써 시착 결과의 정확성을 더 높일 수 있다!

1. **포즈 키포인트 정규화**

Jp와 Jg의 포즈 키포인트를 [0, 1] 범위로 정규화하여 네트워크 입력으로 사용

1. **시착 조건 입력 생성**

ctryon = (Ia, Jp, Ic, Jg)로 정의된 시착 조건 입력 생성