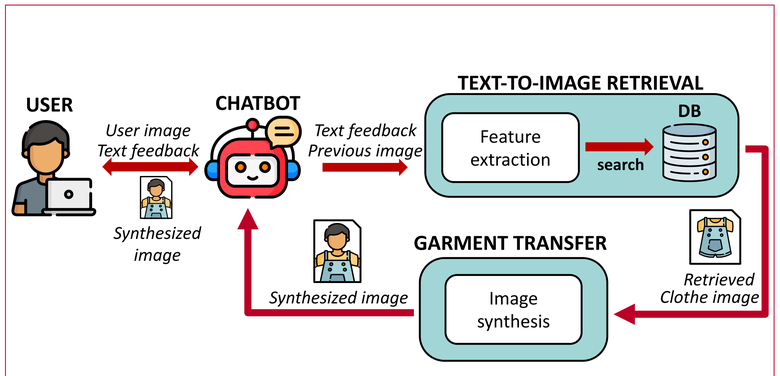
TRY-CIGAR: TRY-on System with Conversational Image Retrieval and GARment Transfer

Team 14
Dain Kim, Jooeun Son, Chanyang Seo, Deokhyung Kang

Intro

- Our project developed a new framework that combines **an image retrieval system** and **a garment transfer system**.
- Users will input desired **attributes of the garment** they are looking for, after which the system will output try-on images of matching items in the database.
- The query and image generation will be done in an **iterative fashion** until the user is satisfied with the results of the query.
- The user may provide multiple attributes, where each providing will yield a try-on image that is compatible with previously input attributes.

System Architecture



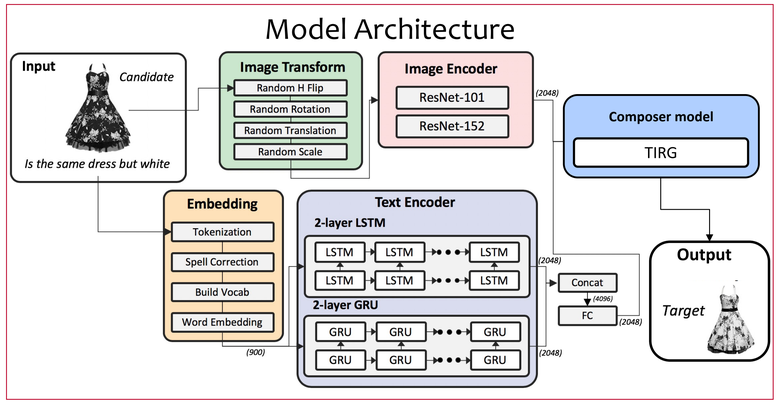
Pipeline

1. **Initial Phase**
   - A user selects a category and attributes of a garment.
2. **Image Retrieval**
   - The system shows top-10 garment image resulted from the image retrieval system.
   - A user chooses a cloth he/she wants to try on.
3. **Garment Transfer**
   - The system shows a try-on image.
4. **Feedback Phase**
   - The user can either continue to search for different cloth by providing a text feedback or stop searching.

Image Retrieval

Baseline Model

- The 2nd place model in Fashion IQ Challenge.
- To reduce the number of parameters of the overall system, we used a single model, TIRG which is a part of ensemble models.

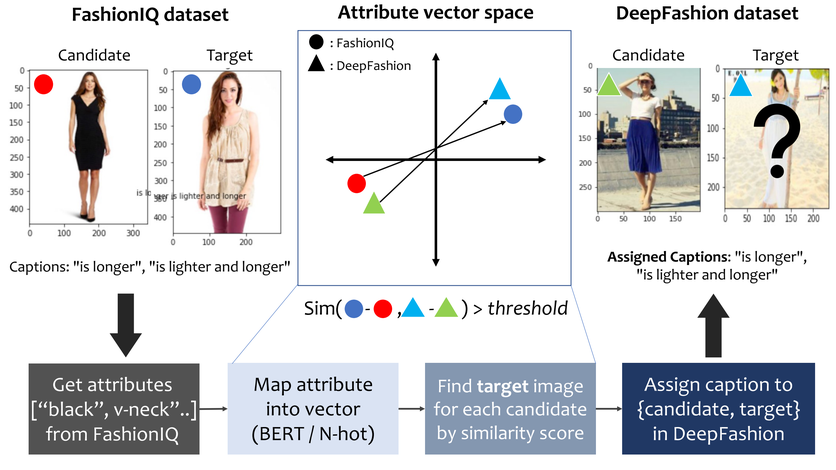


<Baseline Evaluation>

Recall@10	Recall@50
0.226	0.485

Modifications: Data Generation

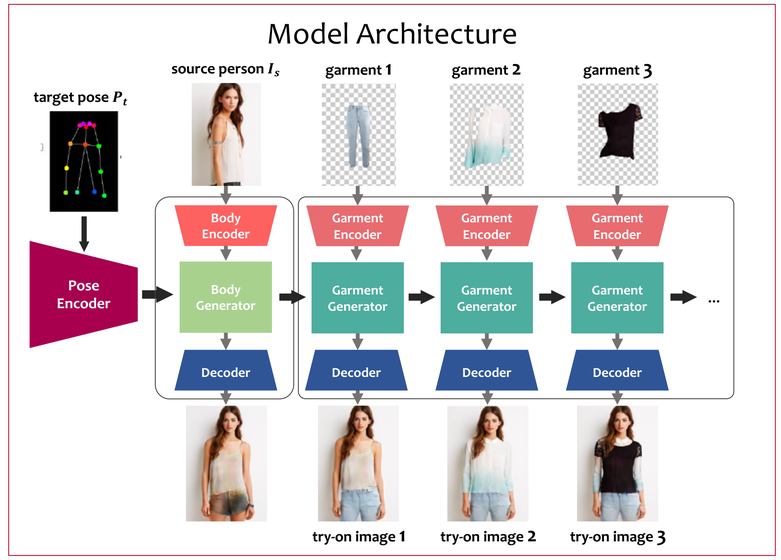
Annotated caption data for lower-body clothes does not exist in DeepFashion. To mitigate this issue, we **generated captions** using the difference between the attributes of a candidate and a target image.



Garment Transfer

Baseline Model: Dressing in Order (2021)

- Network for pose transfer and virtual try-on.

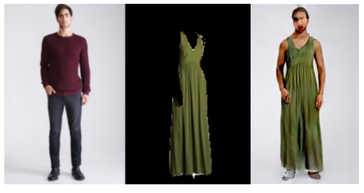


- **Input:** Source image, target pose, garment images.
- **Output:** Recursive generation of VTON Results.

Modifications

1. Integration with IR system-generated outputs
   Automatic pre-processing of IR system outputs before inputting to the network (Pose extraction, Segmentation)
2. Added support for dress category

Automatic Pose and Parsing Extraction



Transferring Dress Category

Limitations

1. Identity Distortion
   Baseline tackles the task of pose transfer
   → Retraining to not change pose
2. Failure cases with complex background images.
   → Trained on white-background images, possibly solved by background augmentation