**Project Report: Multi-Stage Fashion Detection and 3D Visualization System**

**Title**

**Multi-Stage Pipeline for Jewellery and Dress Detection with 3D Neural Rendering Using YOLOv8 and NVIDIA Instant-NGP**

**1. Introduction**

This project presents a multi-stage, computer vision-based pipeline for detecting jewellery and dresses in videos and converting them into a 3D view using NVIDIA's Instant-NGP. The system is designed for fashion and virtual try-on applications, providing users with a dynamic and interactive visualization of how jewellery and clothing appear in real-life scenarios.

**2. Pipeline Overview**

**Step-by-Step Workflow:**

1. **Video Input**: The system starts with a video input.
2. **Jewellery Detection with YOLOv8**:
   * The video is fed into a YOLOv8 model trained on a custom jewellery dataset sourced from Roboflow.
   * All frames are scanned and jewellery is detected and labelled.
   * Output: A video showing bounding boxes around detected jewellery.
3. **Dress Detection with YOLOv8**:
   * The jewellery-detected video is then processed by another YOLOv8 model trained on women's dresses.
   * Output: A video with both jewellery and dresses labelled in each frame.
4. **3D Transformation via Instant-NGP**:
   * The final detected video is converted into a dataset of frames.
   * These frames are fed into NVIDIA Instant-NGP to create a 3D NeRF-style visualization.
   * Output: A 3D rendered configuration viewable from multiple angles.

**3. Tools and Technologies**

|  |  |
| --- | --- |
| **Component** | **Tool/Framework** |
| Object Detection | YOLOv8 (Ultralytics) |
| Data Labelling | Roboflow |
| 3D Rendering | NVIDIA Instant-NGP |
| Video Processing | OpenCV, FFmpeg |
| Environment | Kaggle (YOLO training), VS Code (NGP) |

**4. Dataset Details**

**Jewellery Dataset:**

* Source: Roboflow
* Classes: Bracelets, Brooches, Belt, Earring, Maangtika, Necklace, Nose Ring, Ring, Tiara

**Dress Dataset:**

* Source: Custom-labelled Roboflow dataset
* Focused on women's fashion and traditional outfits

**5. Implementation Notes**

* Frame extraction and conversion for Instant-NGP done using FFmpeg and Python scripting.
* Instant-NGP dataset follows the standard structure with images/ and transforms.json.
* Pipeline is fully automated until Instant-NGP. NGP currently requires manual or command-line triggering.

**6. Future Work**

**A. Integration of Stable Diffusion**

* Use Stable Diffusion or similar models to enable **jewellery swap-in/swap-out** on 3D models.
* Goal: Visualize different jewellery or dress combinations for a user without retraining models.

**B. Convert NGP Output to Blender-Compatible Models**

* Use tools to convert NeRF output to Blender mesh or textured models.
* Goal: Allow customization based on user-specific hand shape, skin tone, and size.

**7. Outcome**

* Successfully built a multi-stage detection and visualization system.
* Each component is functional, and models are modular.
* Achieved real-time detection and high-quality 3D rendering from video data.

**8. Conclusion**

This pipeline bridges computer vision, deep learning, and neural rendering to build a robust fashion detection and 3D modeling system. It has applications in AR try-on systems, virtual showrooms, and fashion tech startups aiming for real-time garment and jewellery visualization.

**9.Dataset, Repository, Models – Links**

Jewellery Dataset – <https://universe.roboflow.com/evin-bu0tz/jewel-detection>

Dress Dataset – <https://universe.roboflow.com/ai-u6f2r/clothes-3i7n1>

Nvidia Instant-NGP – <https://github.com/NVlabs/instant-ngp>