

Data Generation for Training and Testing Machine Learning Models

Data Generation Techniques and Their Advantages

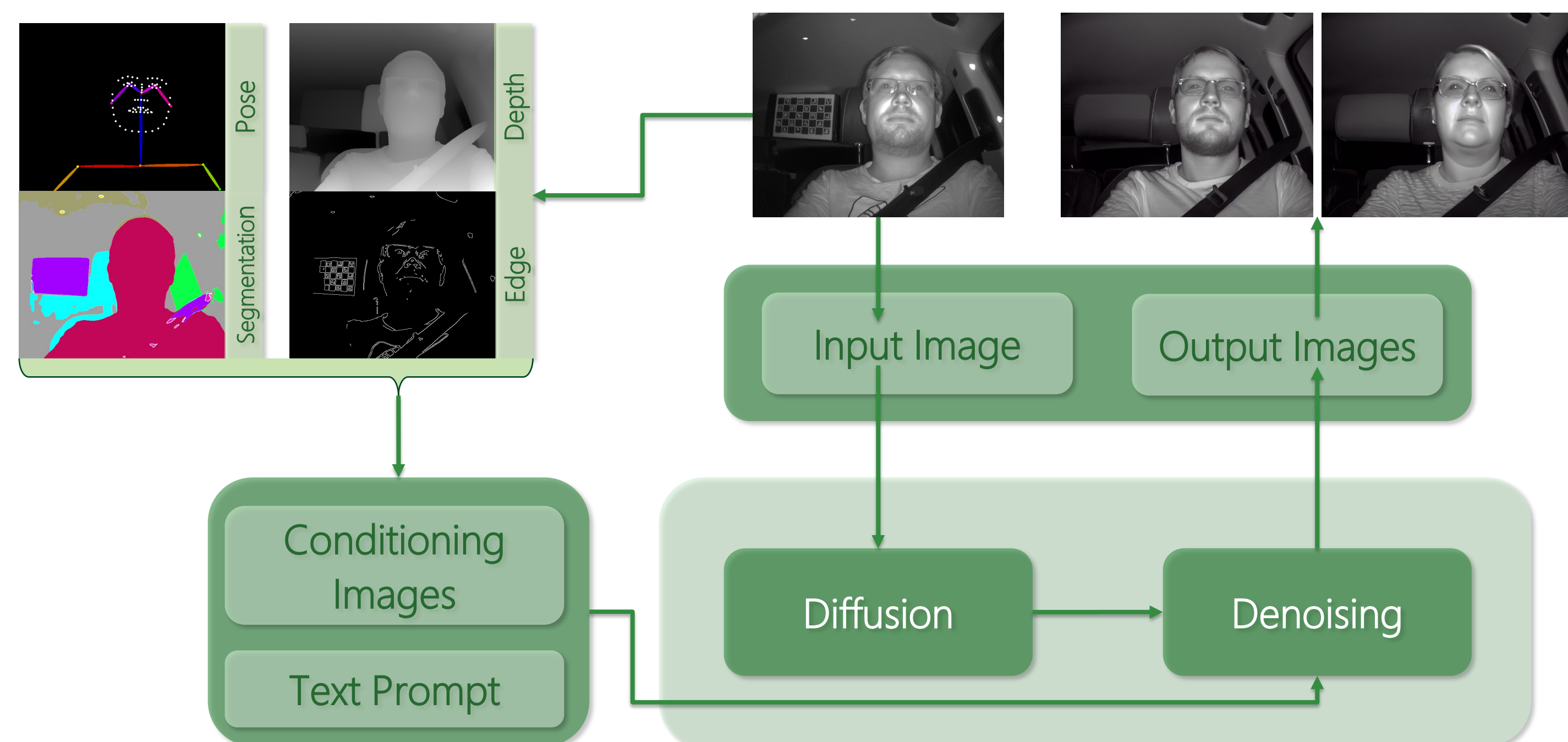
Different approaches are being used to generate synthetic data for different purposes:

- **Image Generation:** Allows the generation of a single frame containing people and objects in the needed modality (RGB or NIR). The person can belong to any ethnicity, age or sex and can be in any pose.
- **Animation Retargeting:** Allows the transfer of micro expressions and accurate head movement from one image of a person to another.
- **Video Generation:** Enables the generation of coherent 5–10 second video clips of people performing pre-designed, predefined movements.

Using these techniques together enables us to generate datasets of images and videos which can be utilized for training and testing ML models that estimate "Driver Readiness" based on "Body Pose" and "Mental State".

Image Generation with a Diffusion Network

By using the information extracted from a real image or a rendered scene as input for a diffusion model, new images with the same structure can be produced.

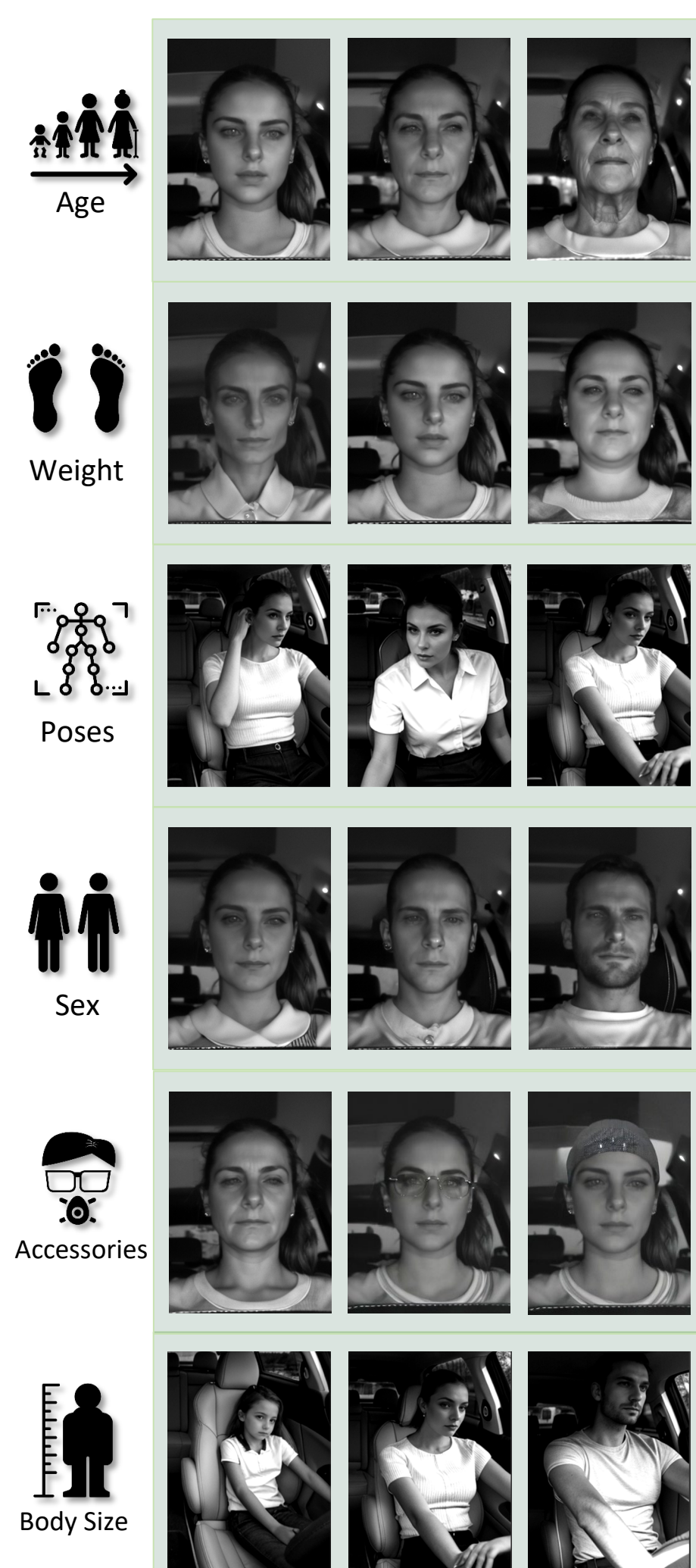


Increasing Data Variety with Image Generation

It is possible to control the generated output by changing the conditioning images and texts. This allows the creation of datasets out of the generated images for different purposes:

- Ethnicity, Sex and Age Estimation
- Weight Estimation
- Discrete Emotion Recognition
- Object Detection
- 2D and 3D Body Pose Estimation

This also allows the introduction of variety into existing datasets especially for underrepresented attributes.

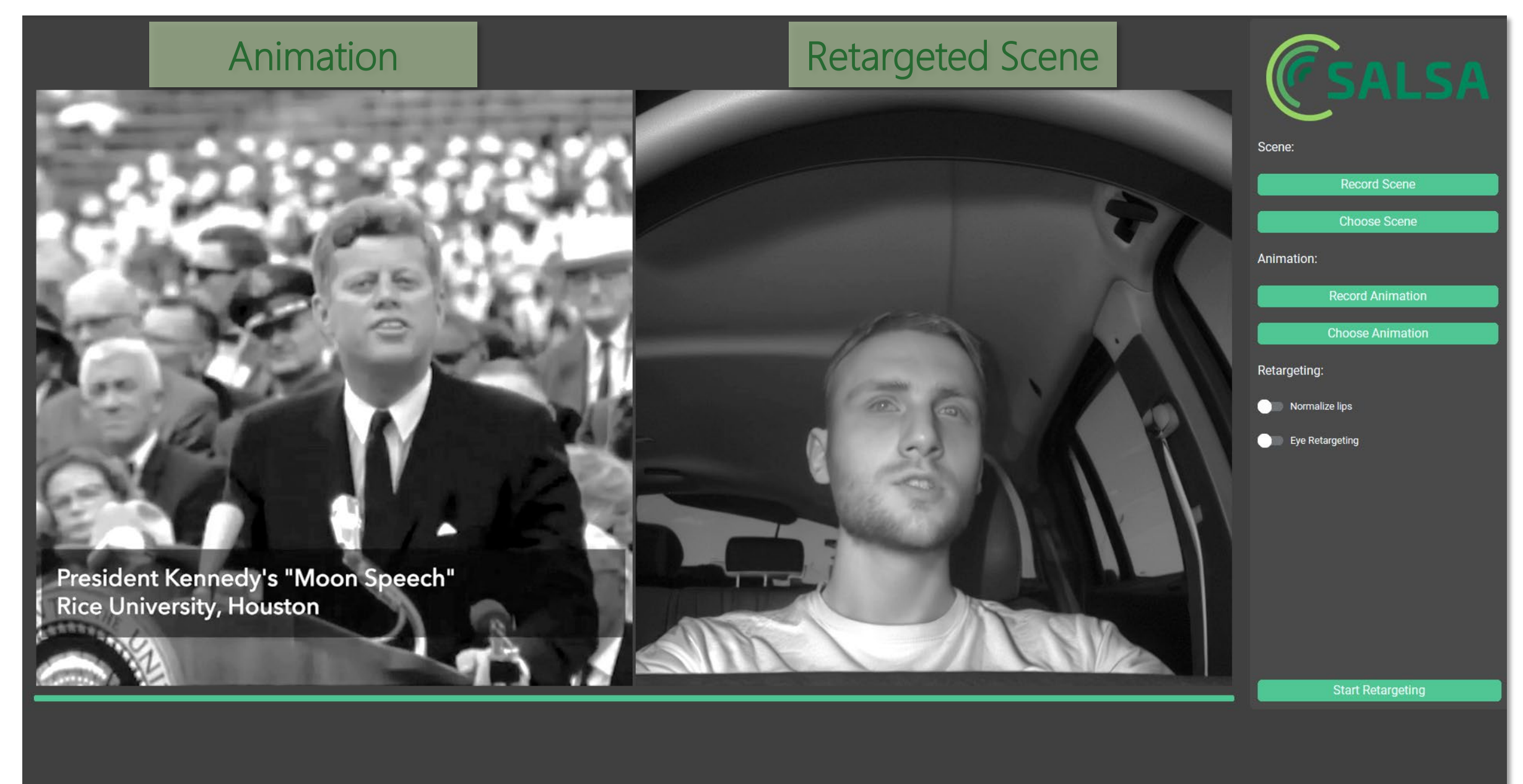


Animation Retargeting

Animation retargeting allows the transfer of

- Facial expression
- Eye opening
- Lip opening
- Head pose

of a person from one image to another. This can be used to transfer motion and expression from real recordings to synthetic scenes, achieving an extra level of realism for synthetic image generation that is not possible when only diffusion models are used.



Video Generation

Unlike Image Generation techniques, Video Generation is specifically designed and optimized for motion and temporal coherence. The network that is being used is much bigger and demanding yet it is still possible to use text, pose and depth conditioning to control the general structure in the generated video. Thus it is possible to generate videos of people moving in a predefined way. This movement can be extracted from real recordings or can also be created using modeling software. Additionally, in order to achieve the best results, Video Generation can be used together with other techniques like Image Generation to provide a first frame or Animation Retargeting to generate more realistic expressions on the generated person's face.

