3D Face Reconstruction and Expression Transfer Using Deep3DFaces

Contents

1 Understanding 3D Face Reconstruction

3D face reconstruction is the process of creating a 3D version of a face from one or more 2D images. This involves several important steps:

1.1 Principles and Components of 3D Face Reconstruction

Key steps include:

- **Landmark Detection**: Identifying key points on the face, such as eyes, nose, and mouth, which helps guide the 3D face's structure.
- **3D Model Fitting**: Adjusting a base 3D face shape (template) to match the detected landmarks. This process is similar to adjusting a clay model to fit the unique features of each face.
- **Texture Mapping**: Taking color and detail from the 2D image and applying it to the 3D model, giving the reconstructed face its original appearance.
- **Pose and Expression Estimation**: Determining the face's orientation and facial expression ensures a natural, realistic model.

The 3D Morphable Model (3DMM) is commonly used, allowing for flexible face shapes and textures by adjusting parameters that recreate almost any face and expression.

1.2 Importance of Accurate 3D Face Models

Accurate 3D models are essential for applications requiring precise facial movement capture, realistic expressions, and adjustments for head angle or lighting.

2 Choosing a 3D Face Reconstruction Model

Our chosen model, **Deep3DFace**, leverages deep learning and 3DMM principles to deliver accurate, high-quality 3D face reconstructions. Built on a **ResNet** backbone, Deep3DFace uses the **Basel Face Model (BFM)** to capture detailed facial shapes, textures, and expressions.

2.1 Deep3DFace

Deep3DFace excels at producing realistic and detailed 3D face models. It uses CNNs to extract features from 2D images and predict 3DMM coefficients, enabling it to accurately reconstruct faces even with challenging lighting, pose, and expression variations. Strengths of Deep3DFace include:

- **High Detail and Realism**: Captures fine facial details and texture.
- **Expression Transfer**: Accurately transfers expressions between faces, making it ideal for avatars and virtual reality.

While Deep3DFace requires powerful hardware for efficient processing, its detailed reconstructions make it the ideal choice for applications demanding realism.

3 Deep3DFaceRecon Pipeline

Deep3DFaceRecon reconstructs a 3D face from a 2D image by estimating shape, texture, pose, and expression using a series of processing steps.

3.1 Input Image Processing

- **Image Normalization**: The 2D input face image is normalized, resized, or transformed to meet the model's input requirements. This ensures consistency in image quality, which improves the accuracy of subsequent steps.
- **Face Alignment**: To standardize face orientation, the input image's face is aligned, ensuring that facial landmarks (e.g., eyes, nose, mouth) match predefined points on a standard template. This alignment enhances the quality of feature extraction.

3.2 2D Feature Extraction

- **Deep Neural Network (ResNet)**: Deep3DFaceRecon uses a convolutional neural network based on the ResNet architecture to extract 2D image features. This model captures essential details like textures, lighting, and shading that are vital for constructing the 3D face.
- **Facial Landmark Detection**: Key landmarks (e.g., eyes, nose, mouth) are detected, providing anchor points for the 3D model, enhancing accuracy, and enabling realistic alignment.

3.3 3D Face Coefficients Estimation

The model predicts various 3D Morphable Model (3DMM) coefficients, including:

- **Shape Coefficients**: Describe individual face geometry.
- **Expression Coefficients**: Capture dynamic facial expressions (e.g., smiling, frowning).
- **Texture Coefficients**: Represent skin color and surface texture.
- **Pose and Lighting Coefficients**: Describe face orientation and light effects.

These coefficients are estimated directly from the 2D image features, defining the 3D face structure.

3.4 3D Face Reconstruction

- **Generate 3D Mesh**: Using the estimated coefficients, a 3D mesh representing the face is created, consisting of vertices in 3D space outlining the face's structure.
- **Texture Mapping**: Texture details (e.g., skin tone) are mapped onto the 3D mesh, adding realism to the face model.
- **Depth and Pose Adjustment**: Based on pose coefficients, the face's depth and orientation are adjusted to match the 2D image's viewing angle.

3.5 Rendering and Post-Processing

- **Rendering**: The reconstructed 3D face is rendered into a visualization that closely resembles the original 2D image.

4 Facial Expression Transfer

Facial expression transfer enables capturing expressions from a source face and applying them to a target face, maintaining the target's identity.

4.1 Capturing and Encoding Facial Expressions

Deep3DFaceRecon captures expressions through **expression coefficients** derived from the 3DMM model. These coefficients quantify the source face's expression, allowing accurate transfer.

4.2 How Expression Transfer Happens

Expression transfer with Deep3DFaceRecon follows this process:

1. **Extracting Expression Coefficients**: The source face expression is captured in coefficients that quantify each facial movement.
2. **Applying Expression Coefficients to the Target**: These coefficients are applied to the target face's 3DMM, blending the target's facial structure with the source expression.
3. **3DMM Reconstruction and Rendering**: The modified coefficients reconstruct the target face with the source expression, maintaining identity but reflecting the source's emotions.

4.3 Deep3DFace-Based Expression Transfer

Deep3DFaceRecon uses a full 3D mesh to transfer expressions:

- **3D Face Mesh Reconstruction**: Constructs 3D meshes of the source and target faces.
- **Expression Warping**: Adjusts the target mesh to replicate the source's expression.

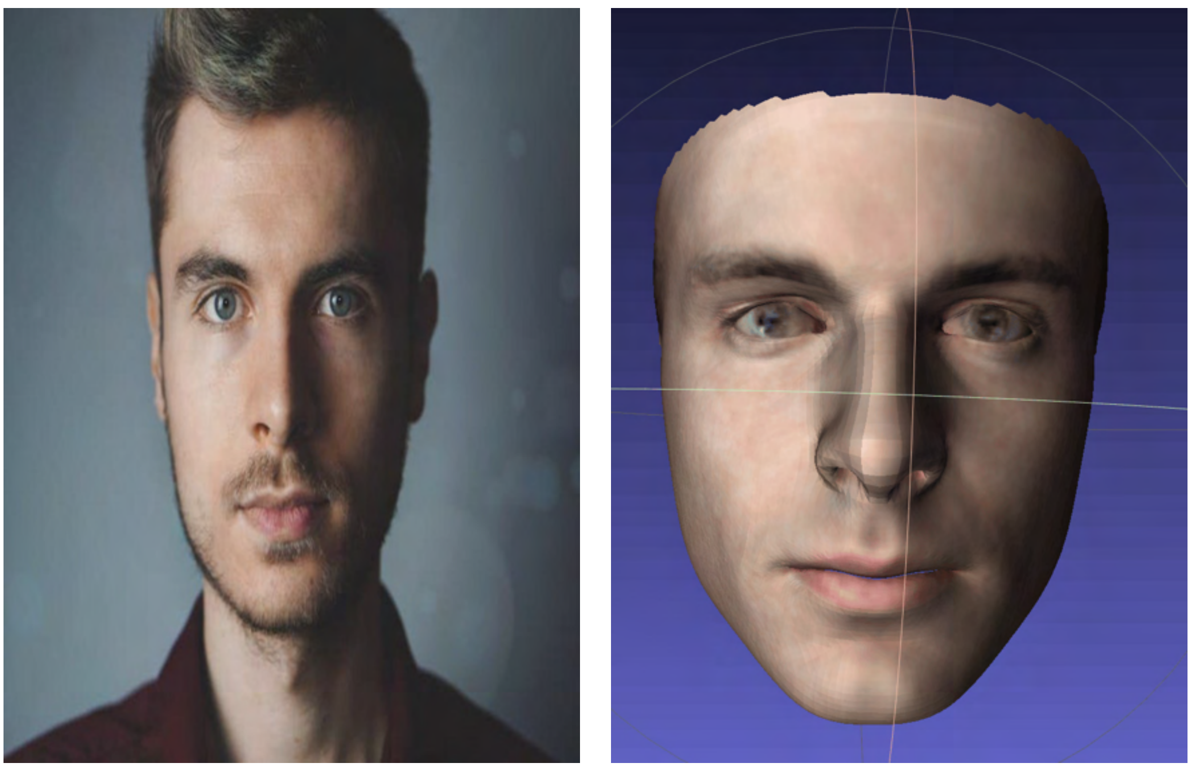
5 Challenges

Several challenges were encountered during the implementation of Deep3DFaceRecon, primarily related to code complexity, dependency management, and computational requirements:

- **Complex Code and Libraries**: Deep3DFaceRecon relies on intricate code with multiple complex libraries, which often lead to compatibility issues. Errors frequently arise due to outdated dependencies or changes in library versions, requiring frequent code updates and adjustments.
- **Computational Requirements and GPU Dependency**: Deep3DFaceRecon is optimized for GPU usage, and running it efficiently on a CPU can be challenging. The model performs best on high-powered GPUs, limiting accessibility for users without this hardware.
- **Lack of Reference and Order for Landmark Detection**: Although Deep3DFaceRecon detects facial landmarks, the order for the five key landmarks needed for face alignment with the input image is not clearly documented, complicating alignment and affecting model accuracy in customized applications.
- **Running and Updates**: Instructions for running and updating the model are provided in the README file.

6 Experiments and Results

6.1 3D Face Reconstruction of Input Images



Original Image 1 | 3D Reconstructed Face 1

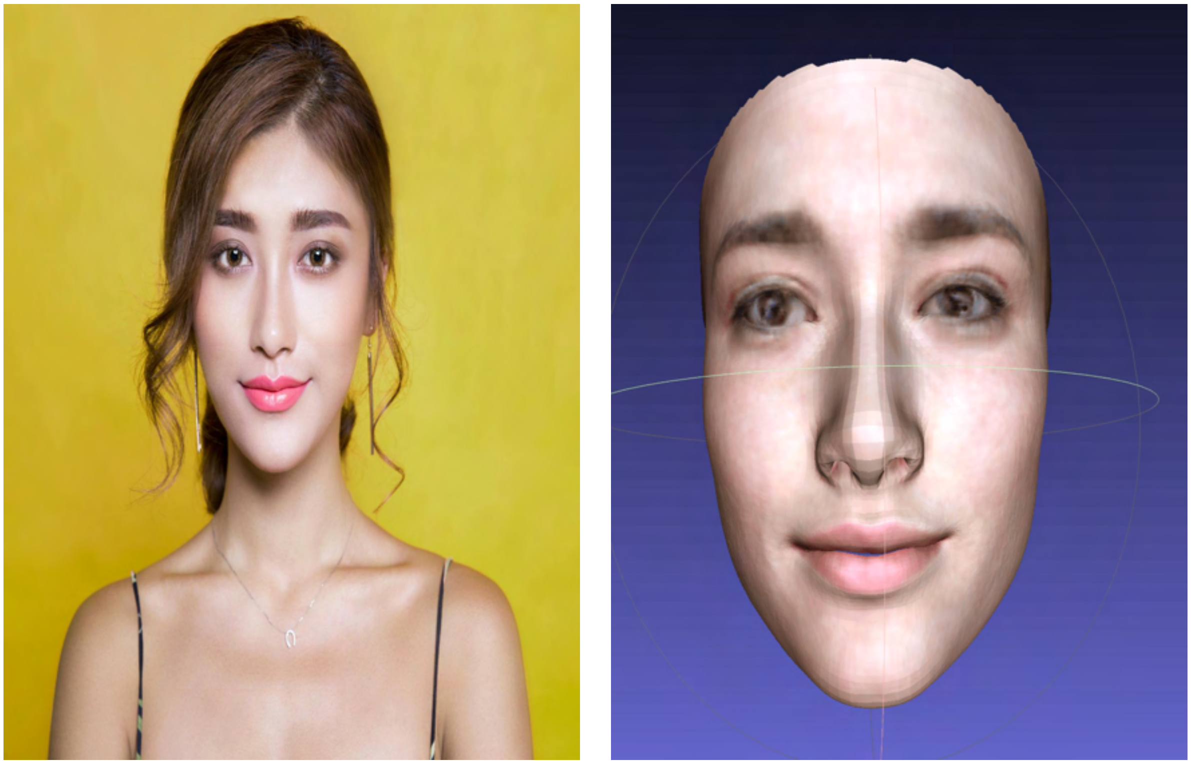
Figure 1: Comparison of original image and it's 3D reconstruction

Original Image 2 3D Reconstructed Face 2

Figure 2: Comparison of original image and it's 3D reconstruction

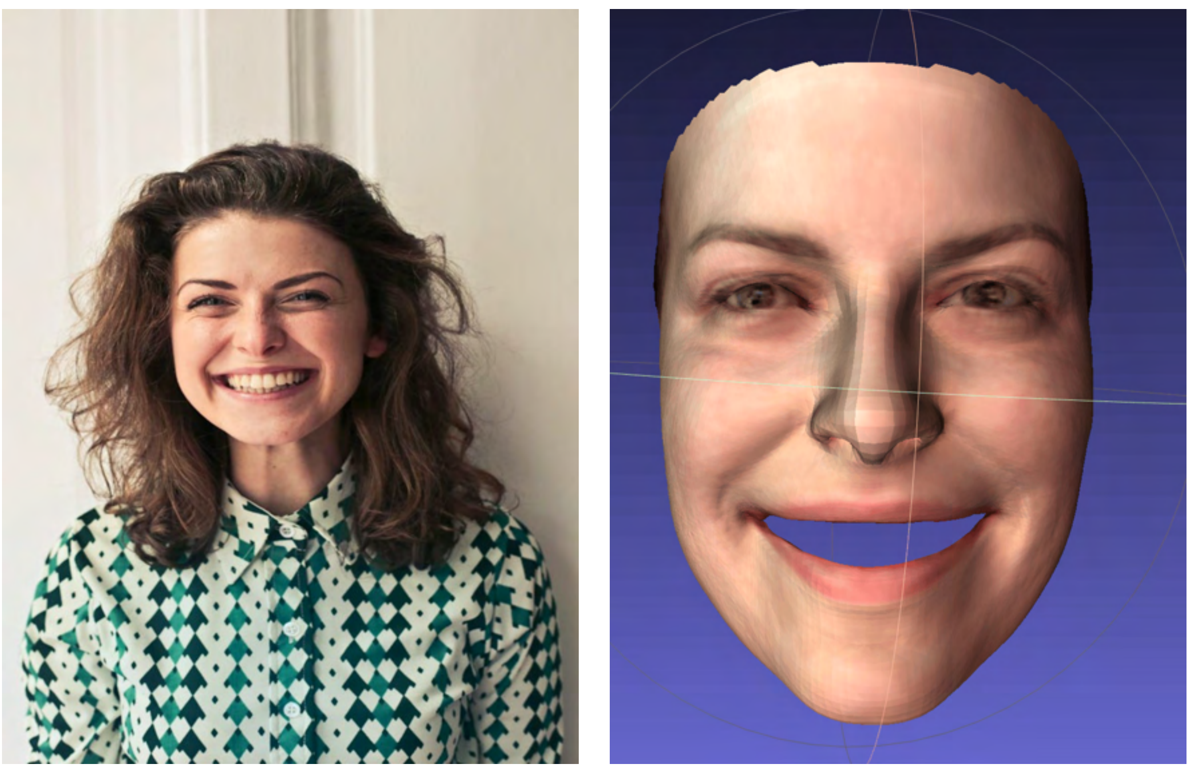
Original Image 3 3D Reconstructed Face 3

Figure 3: Comparison of original image and it's 3D reconstruction



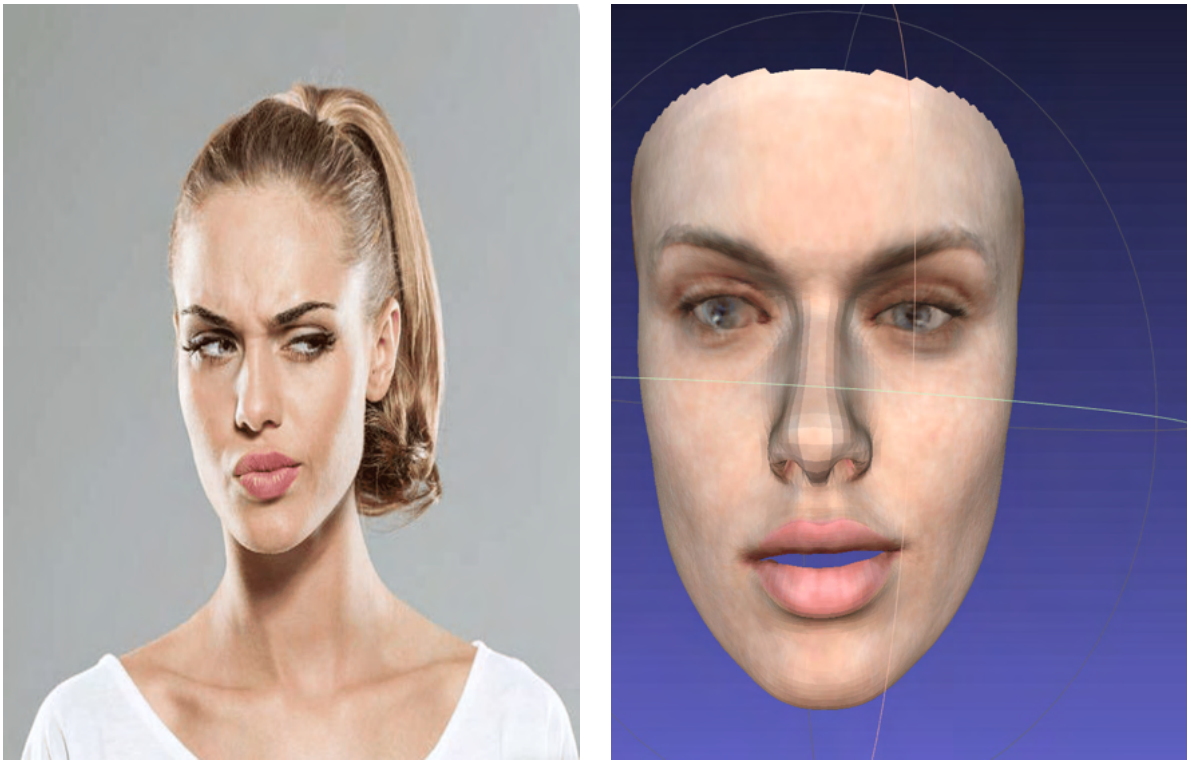
Original Image 4 | 3D Reconstructed Face 4

Figure 4: Comparison of original image and it's 3D reconstruction



Original Image 5 | 3D Reconstructed Face 5

Figure 5: Comparison of original image and it's 3D reconstruction



Original Image 6 | 3D Reconstructed Face 6

Figure 6: Comparison of original image and it's 3D reconstruction

Original Image 7 | 3D Reconstructed Face 7

Figure 7: Comparison of original image and it's 3D reconstruction

Original Image 8 3D Reconstructed Face 8

Figure 8: Comparison of original image and it's 3D reconstruction

Original Image 9 3D Reconstructed Face 9

Figure 9: Comparison of original image and it's 3D reconstruction

6.2 Expression Transfer Test Cases

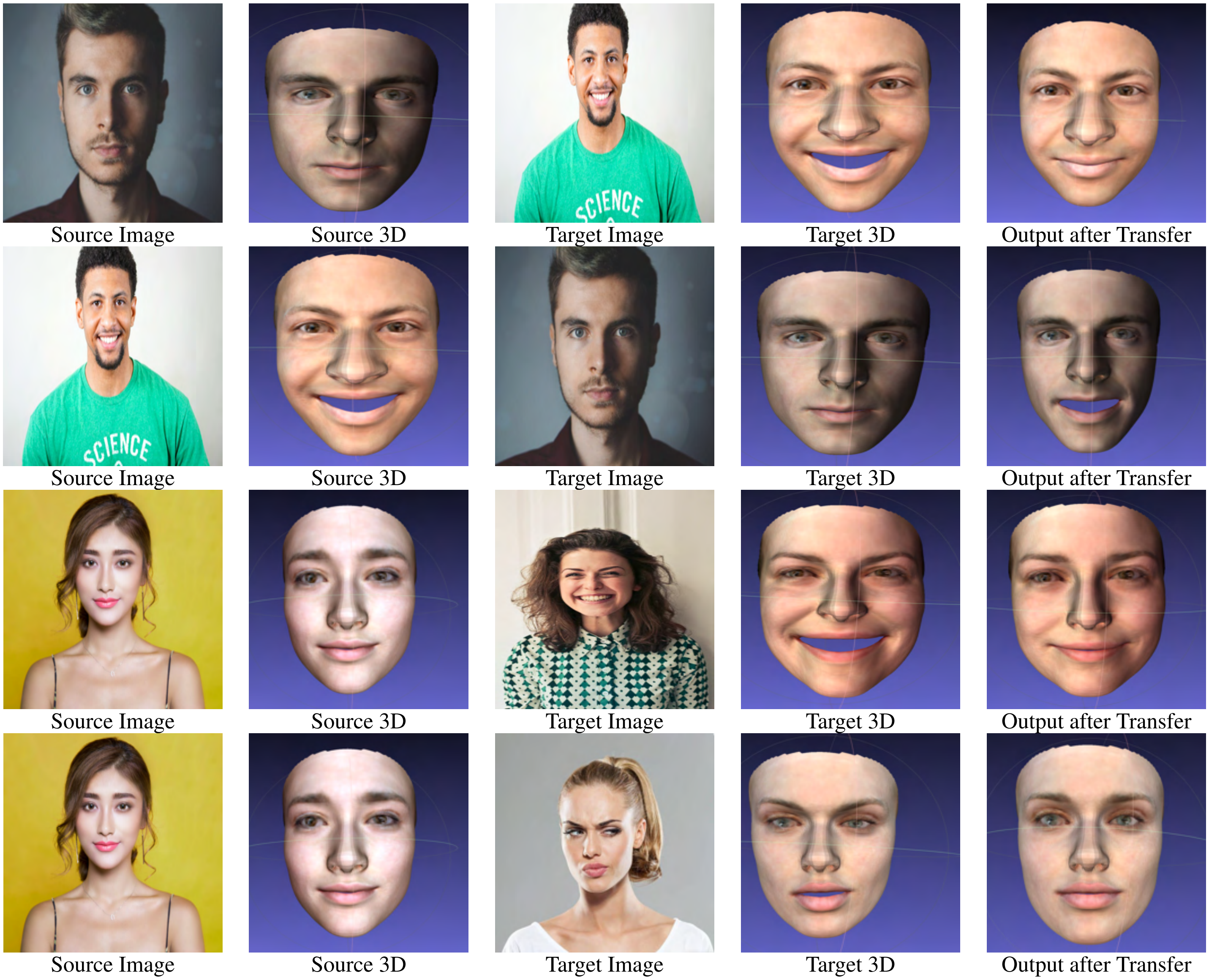


Figure 10: Expression transfer results for first set of test cases

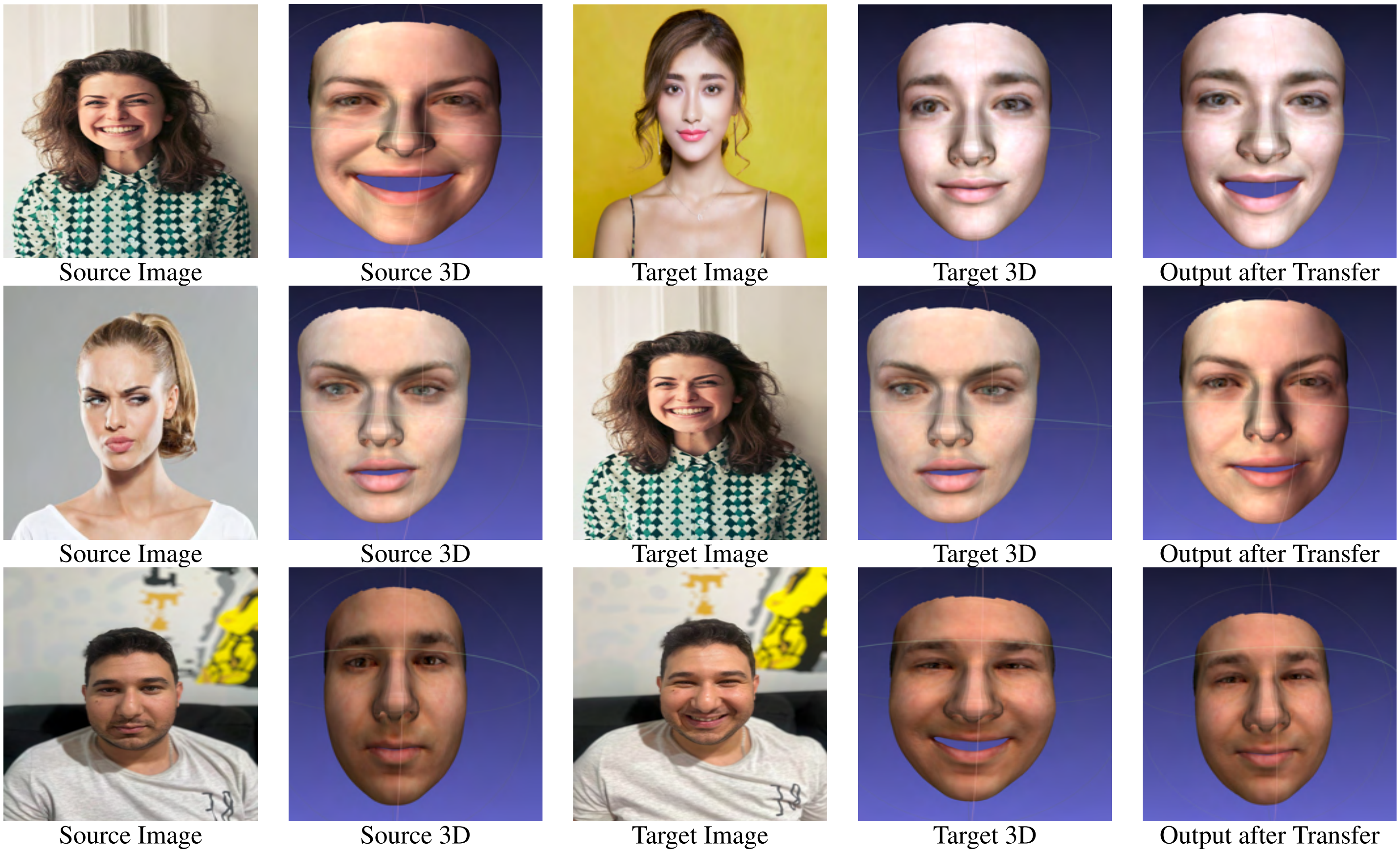


Figure 11: Expression transfer results for second set of test cases

These are screenshots of the 3D expression transfer outputs. For a full 3D visualization, you can download the '.obj' files from the following **Google Drive Link** and view them using MeshLab. A link to download MeshLab can be found in the references section.

7 Conclusion

Deep3DFaceRecon offers a robust solution for 3D face reconstruction and expression transfer, excelling in both realism and detail. By combining the ResNet architecture with the Basel Face Model, it accurately captures facial structures, textures, and expressions. Though it has high computational requirements and some setup complexity, its capabilities make it suitable for applications in VR, animation,digital avatars,and facial expression transfer.

8 References

- **Deep3DFaceRecon GitHub Repository**: This repository provides the official implementation of Deep3DFaceRecon by Microsoft Research, including code, dependencies, and installation instructions for high-fidelity 3D face reconstruction.
- **Accurate 3D Face Reconstruction with Weakly-Supervised Learning**: This paper discusses advancements in 3D face reconstruction through weakly-supervised learning, addressing challenges in reconstructing realistic 3D face models from single 2D images.
- **3D Morphable Model**: Overview of 3DMM, commonly used in 3D face reconstruction, with details on its development and applications in facial modeling.
- **MeshLab**: For a full 3D visualization use meshlab to render the '.obj' files.