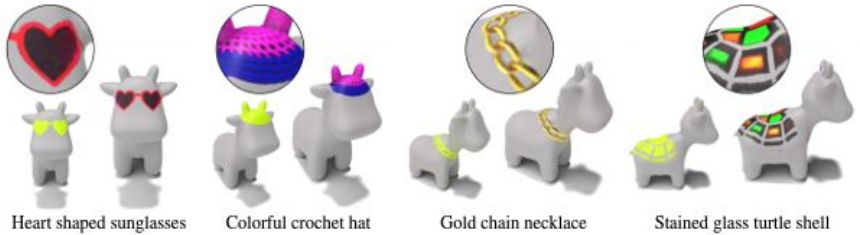
3D Paintbrush: Local Stylization of 3D Shapes with Cascaded Score Distillation

What is the problem?

The problem addressed in the research is the need for localized and detailed text-driven stylization of 3D meshes. While previous methods achieved global text-driven 3D edits, they struggled to make precise, localized modifications on 3D shapes without affecting other areas. Existing methods either worked on voxels or neural radiance fields (NeRFs), but they lacked high-resolution texture generation and fine-grained localization necessary for detailed edits on meshes.

What has been done earlier?

- **Text-to-3D Editing Approaches:** Earlier works mainly focused on global modifications of 3D shapes. These methods used text prompts to supervise the creation of 3D geometry or styles.
- **Mesh Stylization and Neural Radiance Fields (NeRF):** Mesh stylization used neural networks but lacked fine control over local edits. Neural Radiance Fields (NeRF) enabled 3D scene generation with text inputs, but they focused on volumetric rendering, not detailed mesh texturing. Both approaches struggled with precision in local modifications on 3D shapes
- **Challenges with Local Edits**: Earlier works struggled with fine-grained local edits on 3D shapes, relying on coarse segmentation or voxel representations. These methods lacked the precision for sharp boundaries and high-quality textures, resulting in edits that were not detailed or accurate enough for specific localized modifications.



Akanshu Aich,B221008

What are the remaining challenges? What novel solution proposed by the authors to solve the problem?

Remaining challenges:

1. **High-Quality Localized Textures:** Earlier methods struggled to generate detailed textures specifically in localized regions on 3D meshes.
2. **Boundary Precision:** Achieving sharp, accurate boundaries between edited and non-edited areas is difficult, leading to artifacts in previous techniques.
3. **Simultaneous Optimization:** Localization and texture maps were not optimized together, causing either low-resolution textures or inaccurate localizations.
4. **Granularity Control:** Balancing fine details and overall shape coherence remains challenging in localized stylization.

Novel solutions:

1. **Cascaded Score Distillation (CSD):** CSD uses a multi-stage diffusion model to refine textures and localization simultaneously, ensuring high-quality and detailed local edits.
2. **Simultaneous Generation:** CSD optimizes both localization and texture maps together, improving precision and avoiding unwanted overlaps.
3. **Multi-Stage Refinement:** The cascaded diffusion model allows for incremental refinement, enhancing both fine details and boundary accuracy.
4. **Control Over Detail and Coherence:** CSD provides fine control over how granular or globally coherent the textures are, allowing seamless integration into existing 3D workflows

Akanshu Aich,B221008

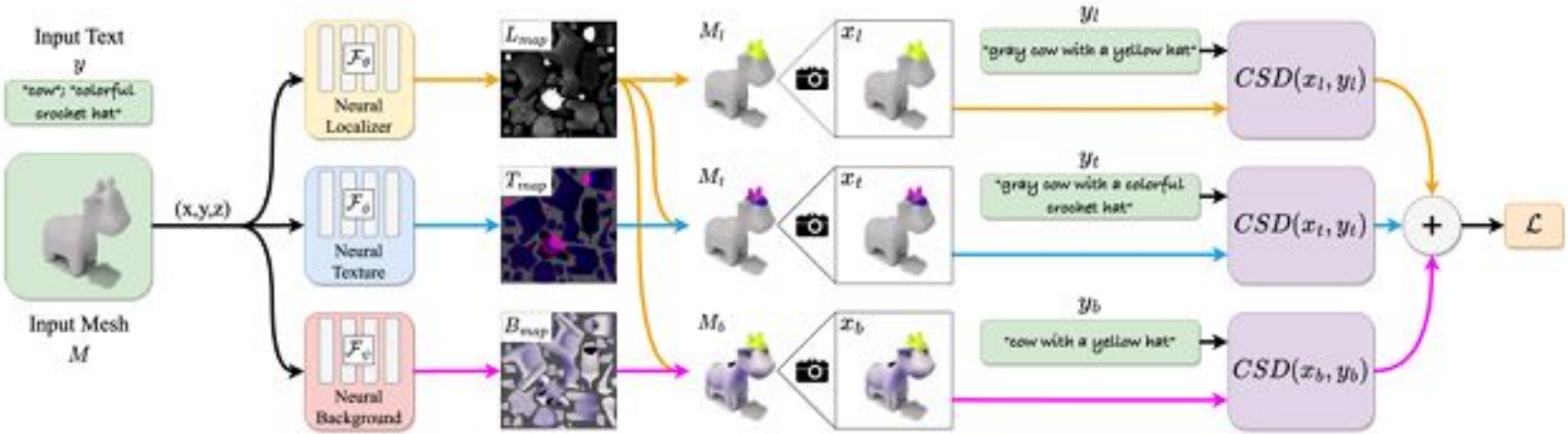


Figure 3. **Overview of 3D Paintbrush**. Each point on the surface of the mesh is passed into three different branches to produce a localization probability, texture map, and background map. We texture three different variants of the same mesh with the localization, texture, and background maps and render them from the same viewpoint. Each image along with the corresponding text condition is used to compute the CSD loss.