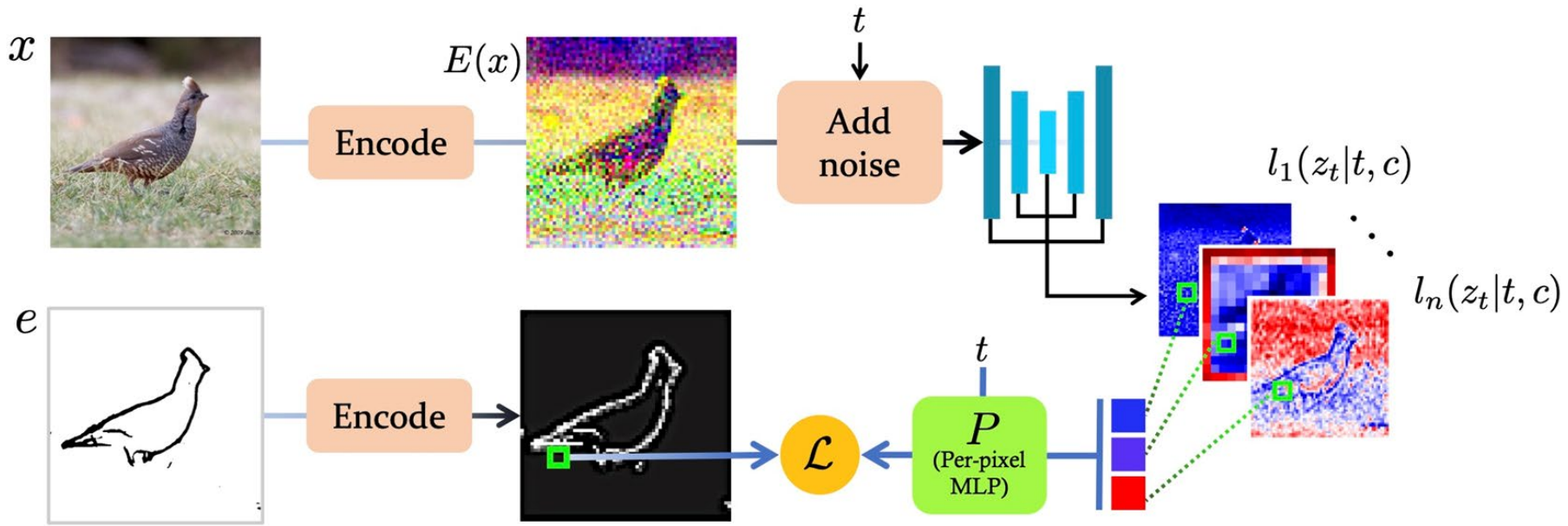
Recap of last week

Sketch-Guided Text-to-Image Diffusion Models



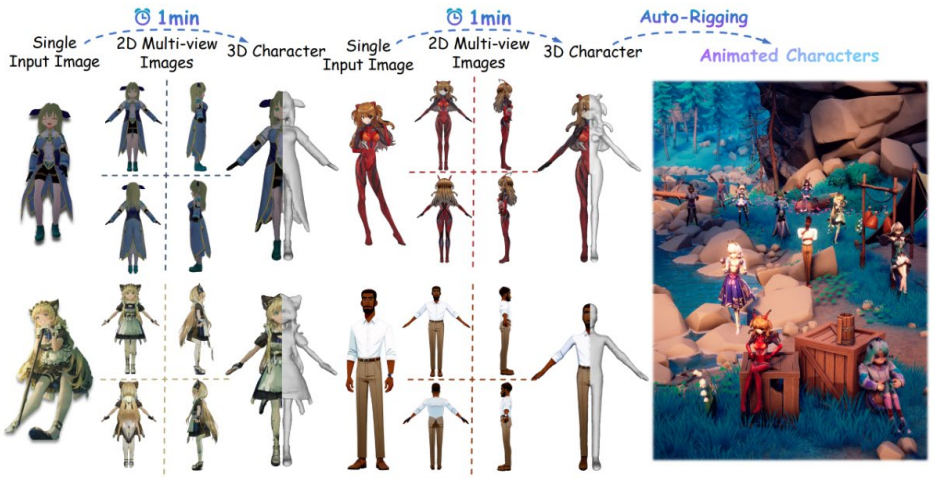
CharacterGen: Efficient 3D Character Generation from Single Images with Multi-View Pose Calibration

SIGGRAPH(TOG) 2024

Hao-Yang Peng[1], Jia-Peng Zhang[1], Meng-Hao Guo[1], Yan-Pei Cao[2], Shi-Min Hu[1†]

[1]Tsinghua University, [2]VAST

[†]Corresponding Author



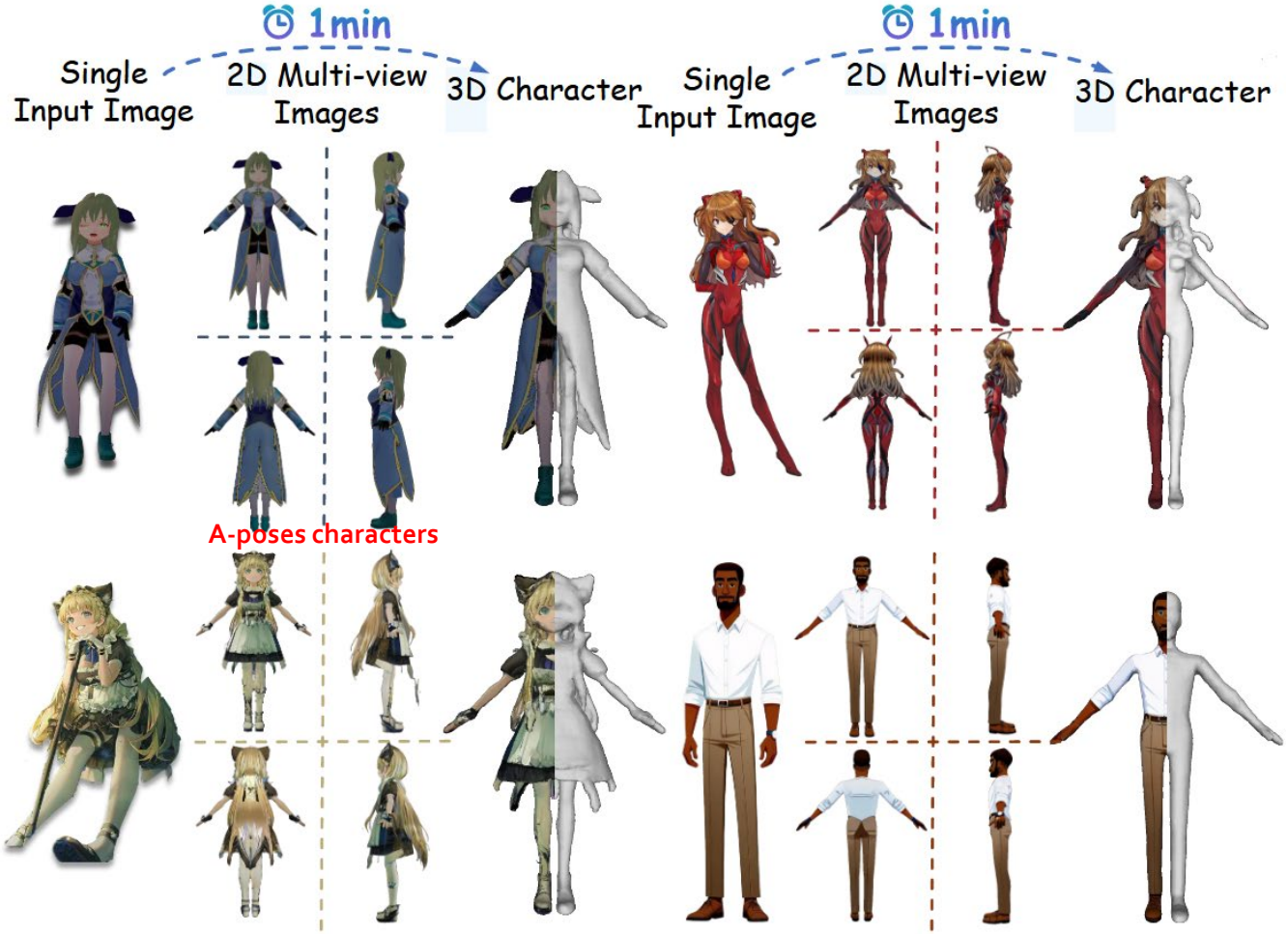
Background

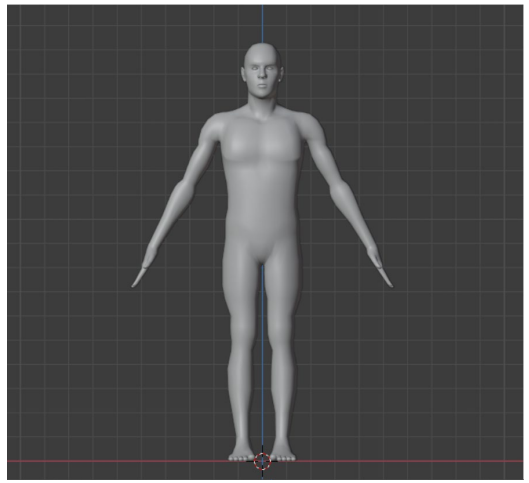
Make-It-3D(ICCV 2023)

Zero 123 (CVPR 2023)

(1) Realistic human style generation
(2) Poses limits

Overall Pipeline

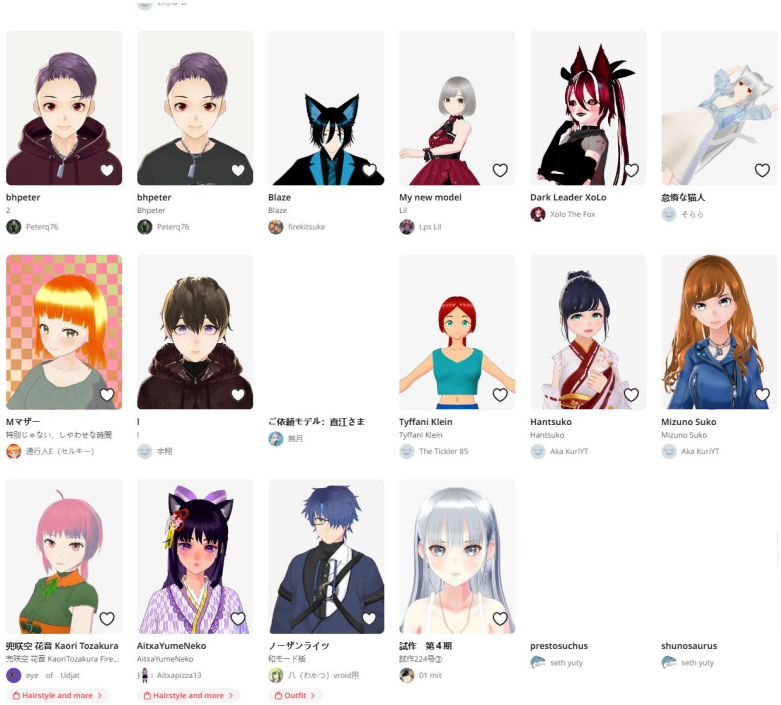




A-poses

Dataset

VRoid Hub 3D character models



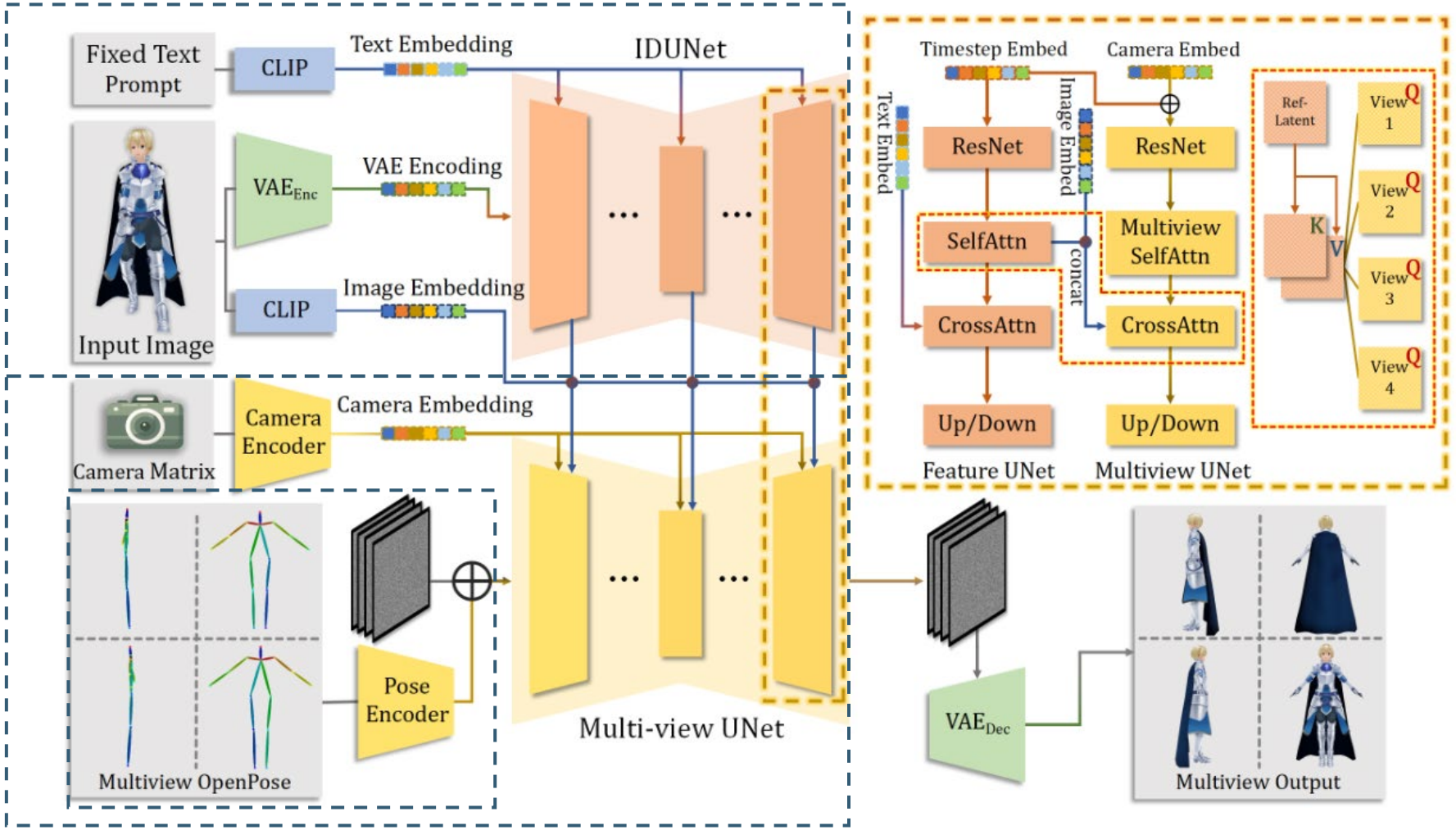
- *.vrm

What is VRM?

VRM can handle `humanoid` `character avatars`.

- The format is **glTF based**, so it's **cross-platform**. It can also be handled by other game engines and the Web.
  - ➡ plugins
- **Provides a standard implementation (UniVRM) for reading and writing VRM**
  - ➡ UniVRM

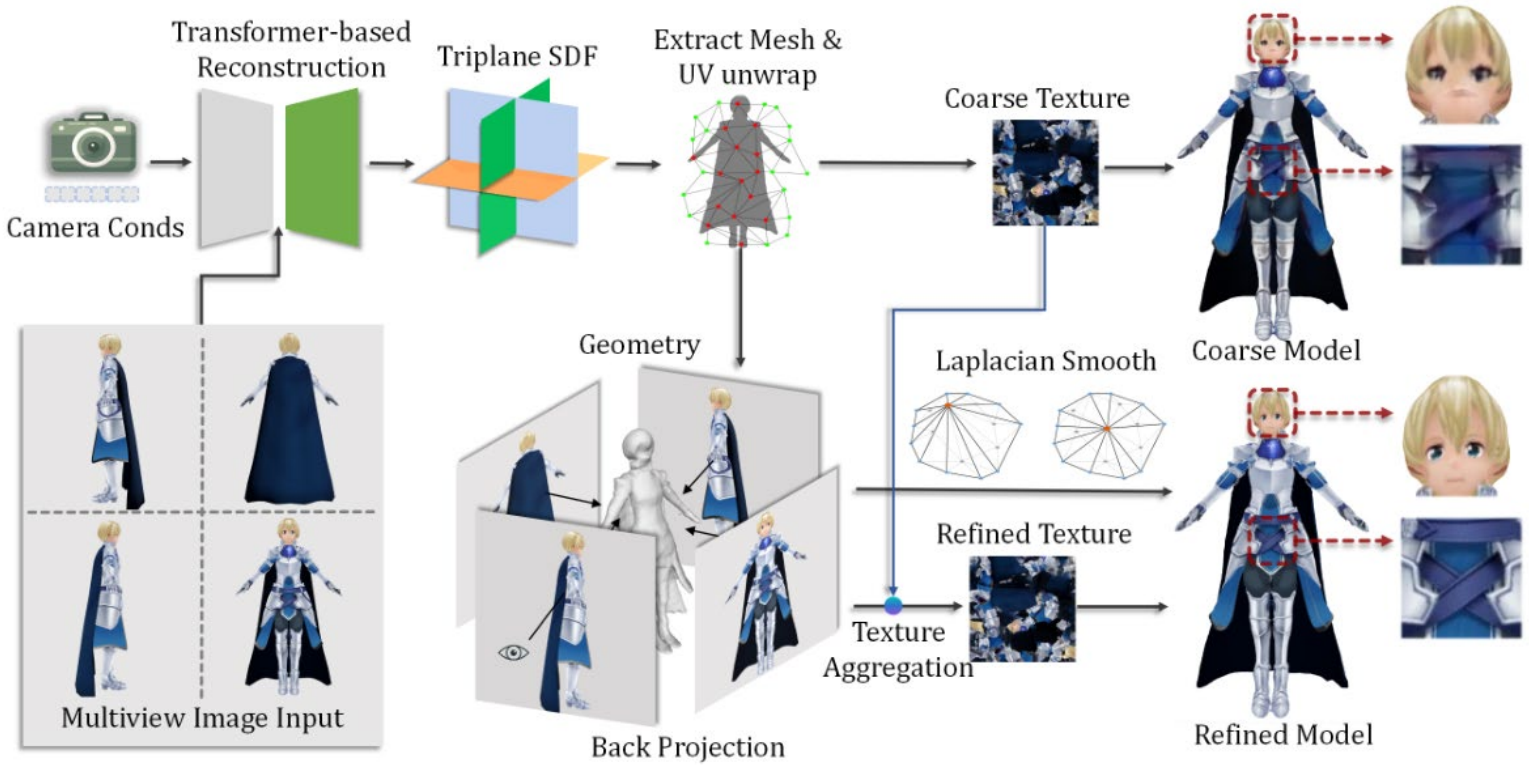
2D multi-view generation



"best quality"

$$f_{\text{Cond}} = \text{concat}(f_{\text{ID}}, f_{\text{CLIP}})$$
$$x_{4v} = \text{Cross\_Attn}(x_{4v}, f_{\text{Cond}})$$

3D character generation: coarse-to-fine



- Loss of Texture Details
- Low-Resolution Input:
- Sparse Input Views
- Silhouette Noise

Experiments

Methods	SSIM↑	LPIPS↓	FID↓	CD↓
CharacterGen(2D)	**0.901**	**0.086**	**0.019**	-
Zero123	0.768	0.224	1.42	-
Zero123(fine-tuned)	0.813	0.175	1.34	-
SyncDreamer	0.807	0.194	0.396	-
SyncDreamer(fine-tuned)	0.822	0.17	0.37	-
IP-Adapter+SDXL	0.845	0.143	0.074	-
CharacterGen(3D)	**0.898**	**0.093**	**0.032**	**0.001**
Magic123	0.873	0.134	0.116	0.0034
ImageDream	0.886	0.11	0.345	0.002

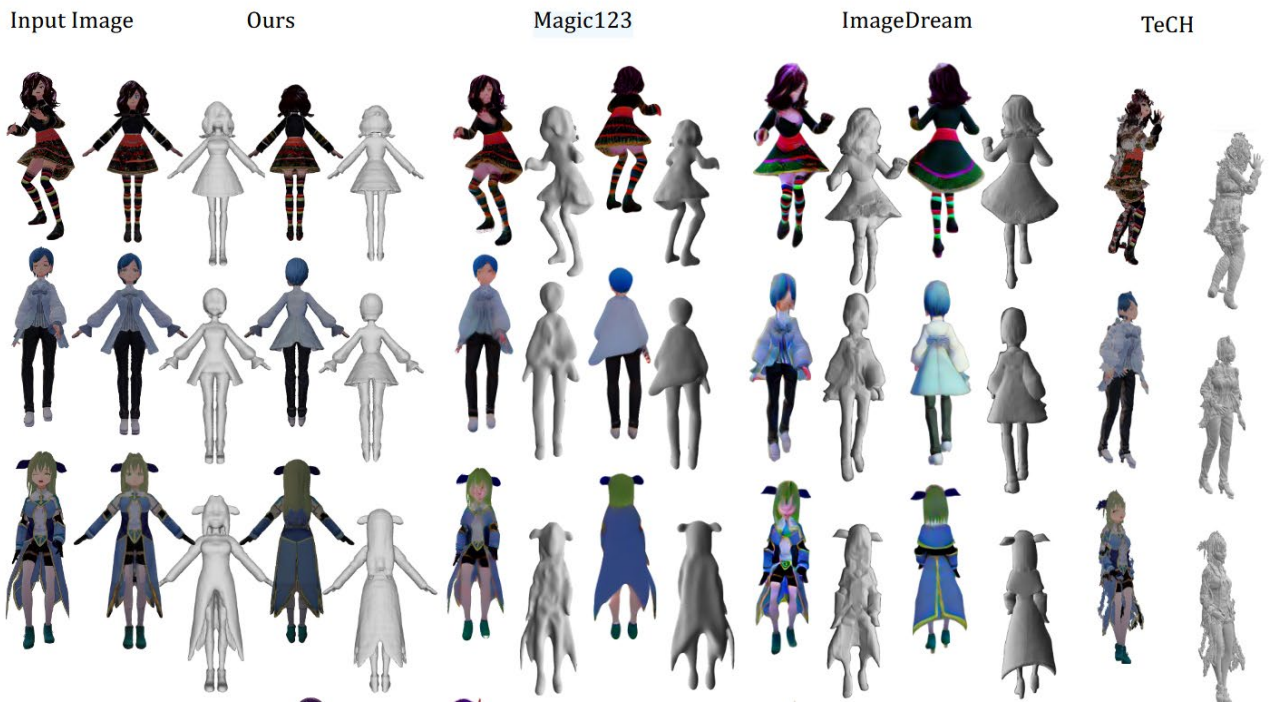
Figure: Effectiveness

Methods	Time
CharacterGen	1min
Magic123 [Qian et al. 2023]	70min
ImageDream [Wang and Shi 2023]	45min
TeCH [Huang et al. 2023b]	270min

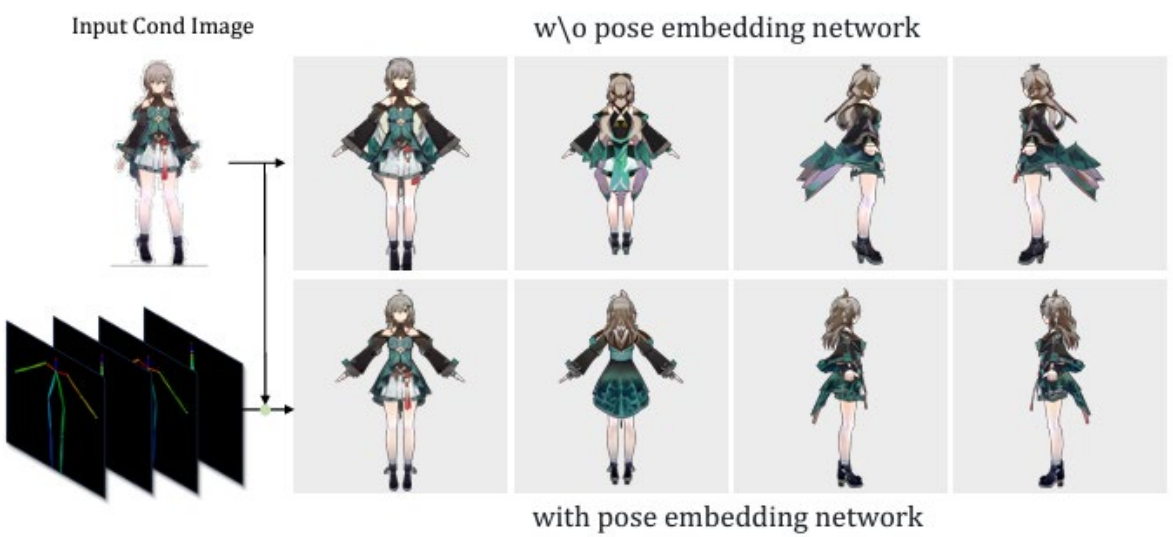
Figure: Time to generate a single 3D character. Models loading time is excluded for all methods.

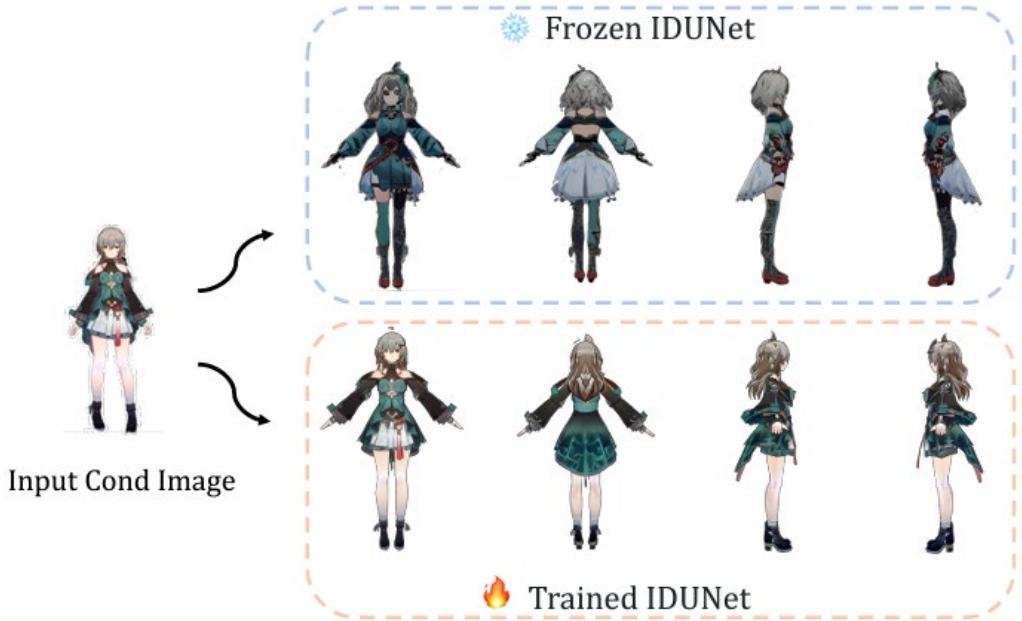
Results

metric	CharacterGen	Zero123(2D)	SyncDreamer(2D)	Magic123(3D)	ImageDream(3D)
2D multi-view style consistency	**85.4%**	10.5%	4.1%	-	-
2D multi-view consistency	**81.0%**	17.1%	1.9%	-	-
3D character geometry quality	**78.6%**	-	-	2.86%	18.6%
3D character texture quality	**87.1%**	-	-	1.9%	11.0%

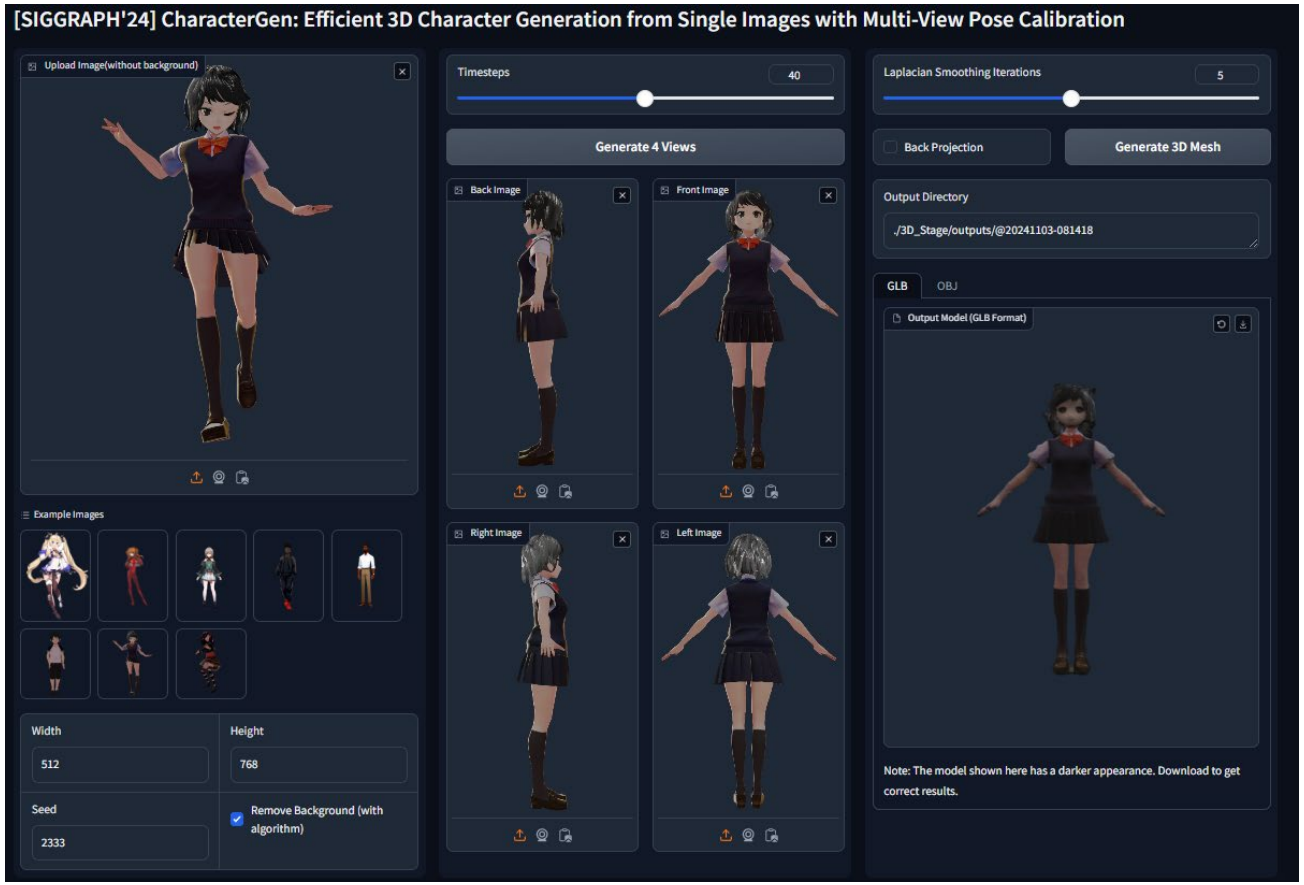


Ablation tests





Conclusions



Quiz

https://forms.gle/WAvmfnvFzMSL48Ni6