



SIGGRAPH 2024

DENVER+ 28 JUL – 1 AUG

DiLightNet: Fine-grained Lighting Control for Diffusion-based Image Generation

Chong Zeng, Yue Dong, Pieter Peers, Youkang Kong, Hongzhi Wu, Xin Tong



Microsoft®
Research
微软亚洲研究院

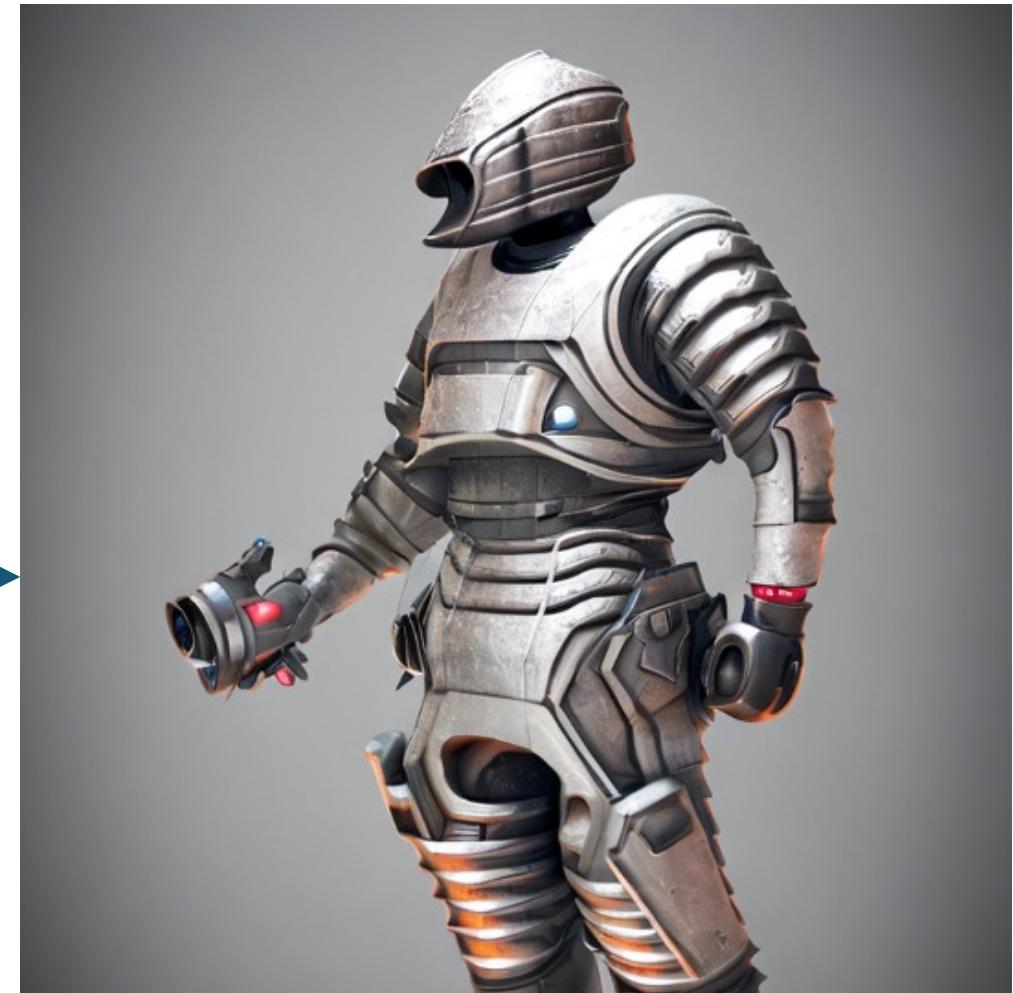
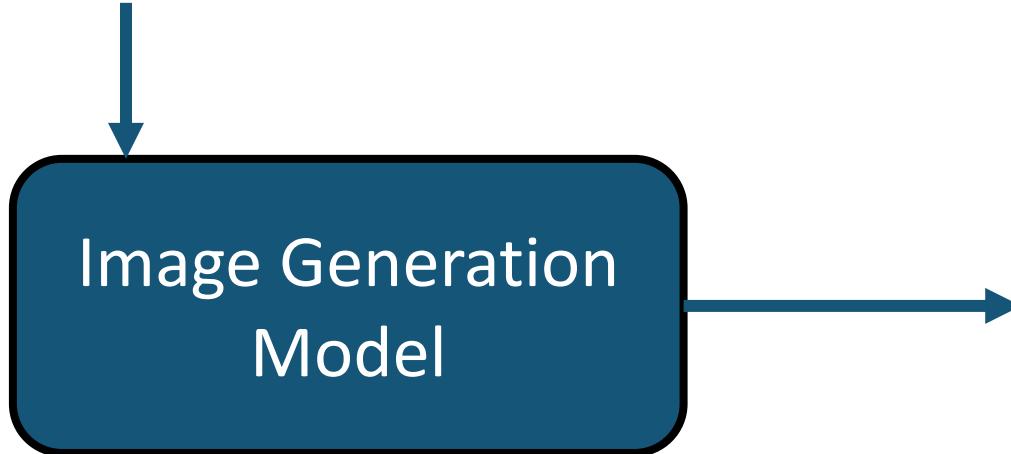

WILLIAM & MARY
CHARTERED 1693



Image Generation with Diffusion Models

Text Prompt

“futuristic soldier with
advanced armor weaponry
and helmet”

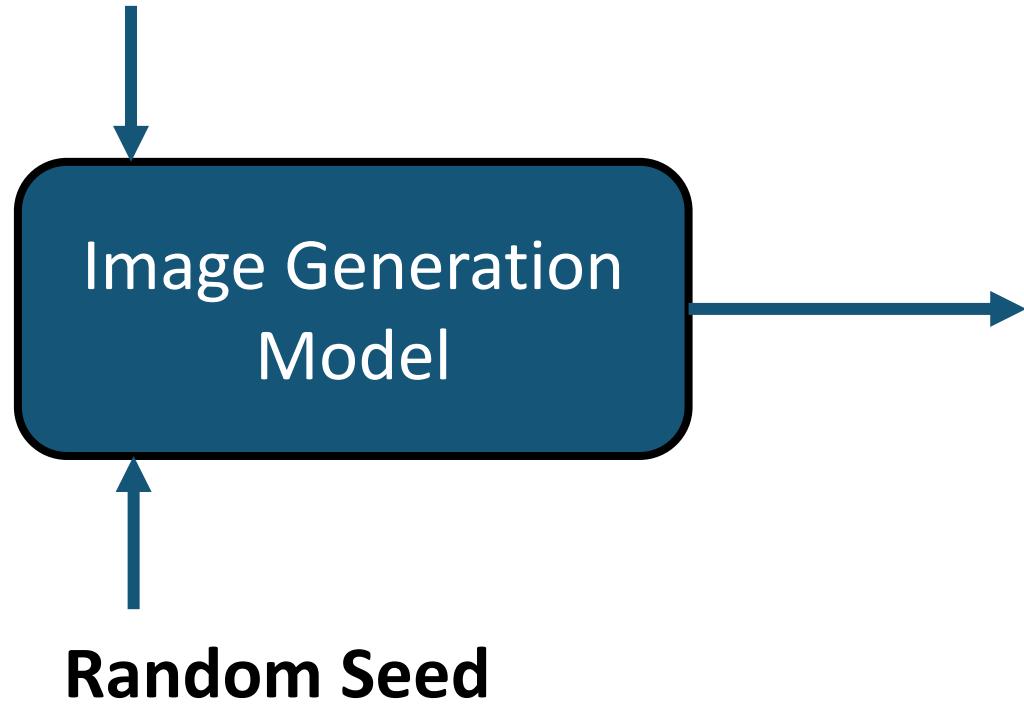


Generated Image

Image Generation with Diffusion Models

Text Prompt

“futuristic soldier with
advanced armor weaponry
and helmet”



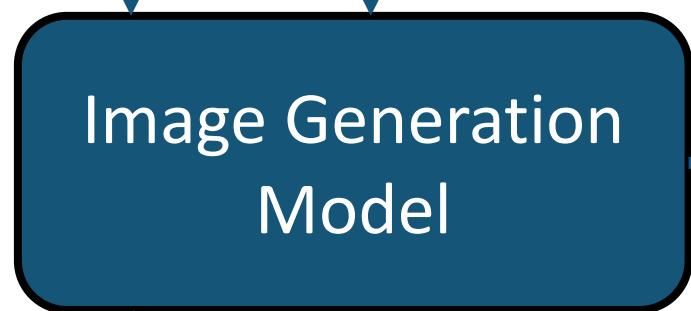
Generated Image

Our Goal: Adding *Fine-grained* Lighting Control

Text Prompt

“futuristic soldier with advanced armor weaponry and helmet”

**Fine-grained
Lighting
Condition**



Generated Image Under **Target Lighting**

“Fine-grained” means arbitrary lighting representation

I want to generate an image
under the lighting of...



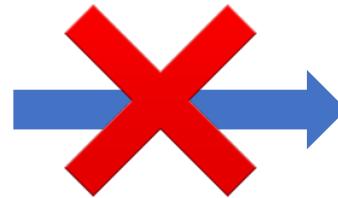
Challenge: Handle Arbitrary Lighting Representation

a colorful indoor
cathedral
environment

a single near-
field **point light**

an **area light**
placed at (1.0,
1.0, 1.5)

“Fine-grained” means keeping image content consistent



This is not what I want...
I tried hard getting the image content!
I want to keep it **fixed**...

“Fine-grained” means keeping image content consistent



Challenge:

Follow Both *Content and Lighting*



That's what I want!

“Fine-grained” includes *Global and Local* Shading Effects

Global Shading Effects

Large Highlights

Complex Shadows

Defined by
Large-scale
Geometry!



“Fine-grained” includes *Global and Local* Shading Effects

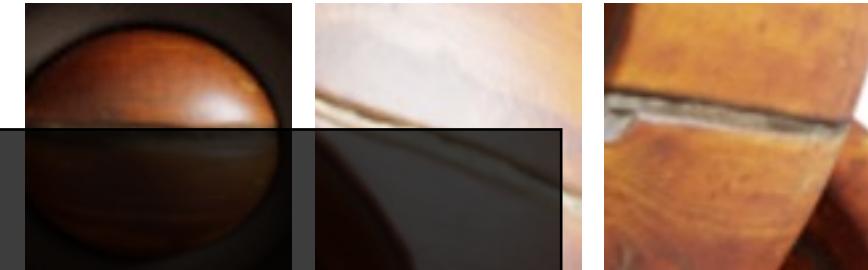
Global Shading Effects

Large Highlights

Complex Shadows

Defined by
Large-scale
Geometry!

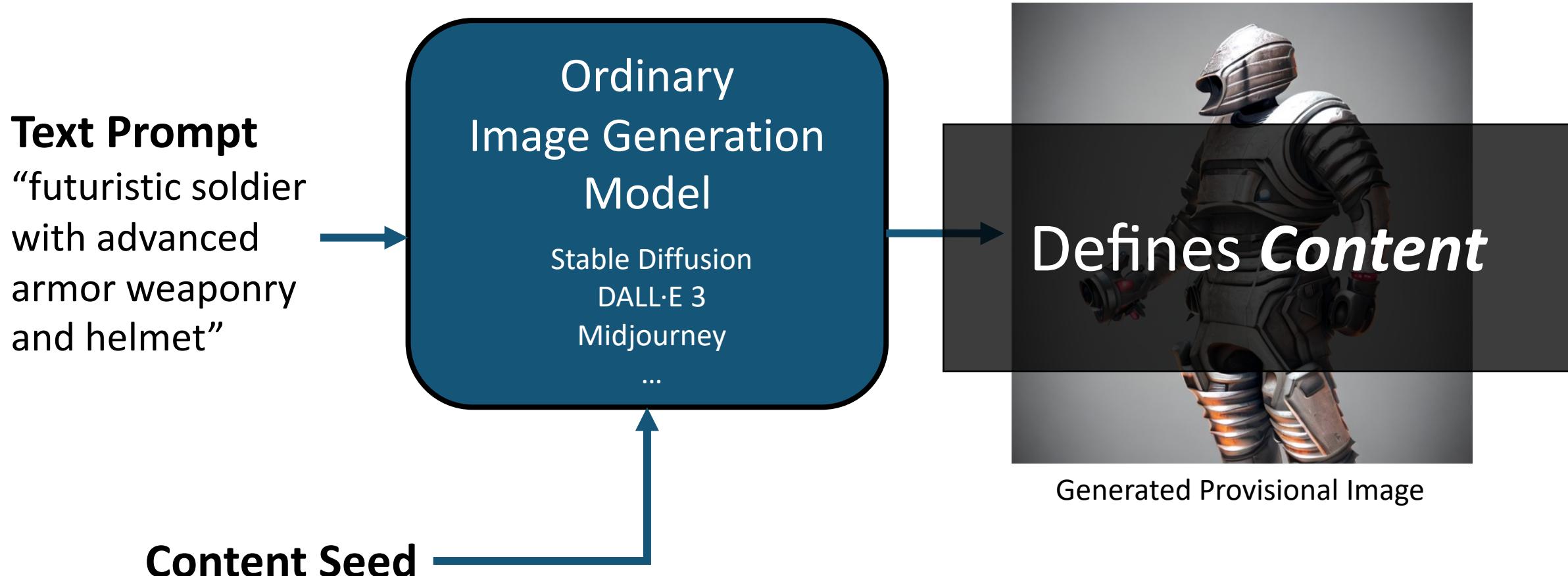
Local Shading Effects



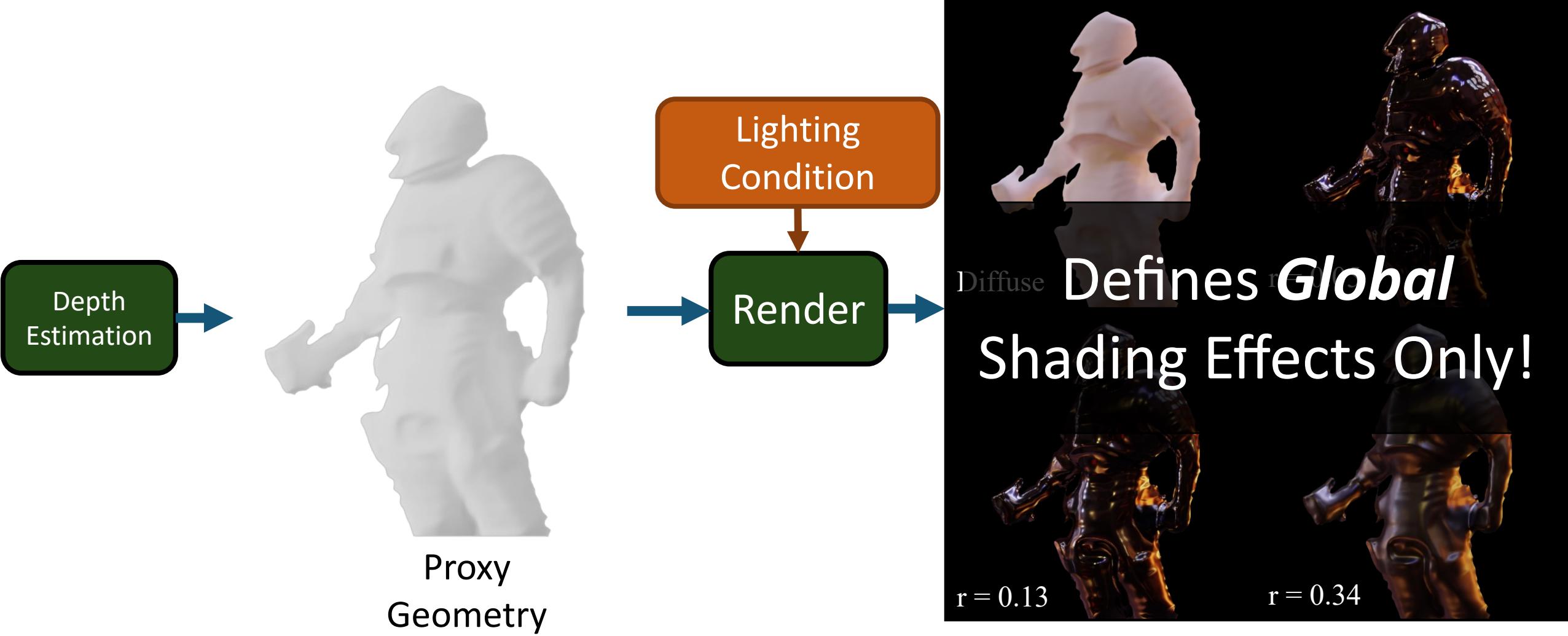
Challenge:
Follow Both *Local and Global* Shading
Details on
Local Details!



Our Solution: Separate Content Generation Stage



Inject Lighting Condition as Radiance Hints



Regard Final Shading Effects as Generative Problem

Material-Lighting
Ambiguity



Provisional Image
Formulate as *generative* problem!

Geometric Detail
Uncertainty

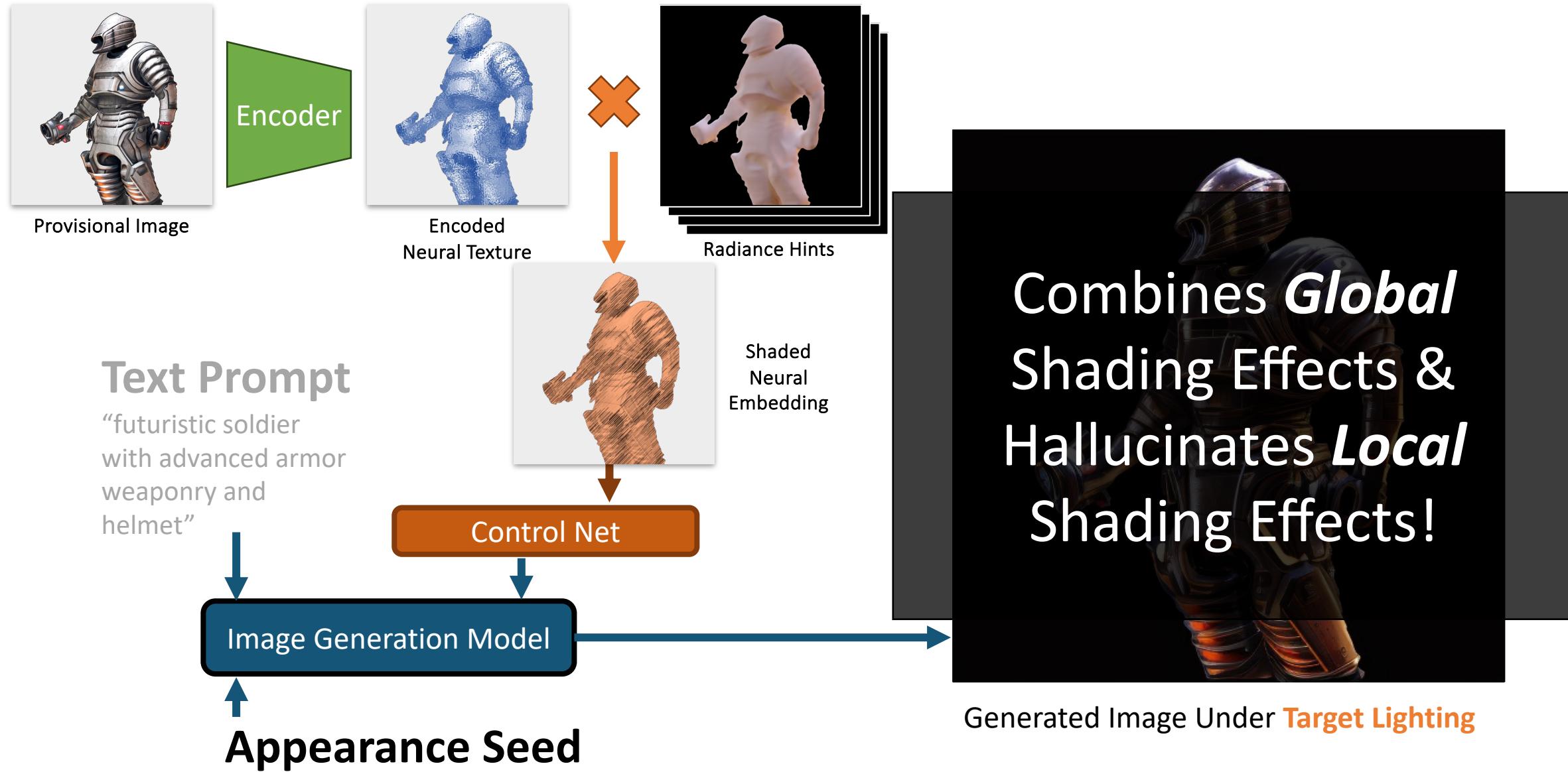


Radiance Hints

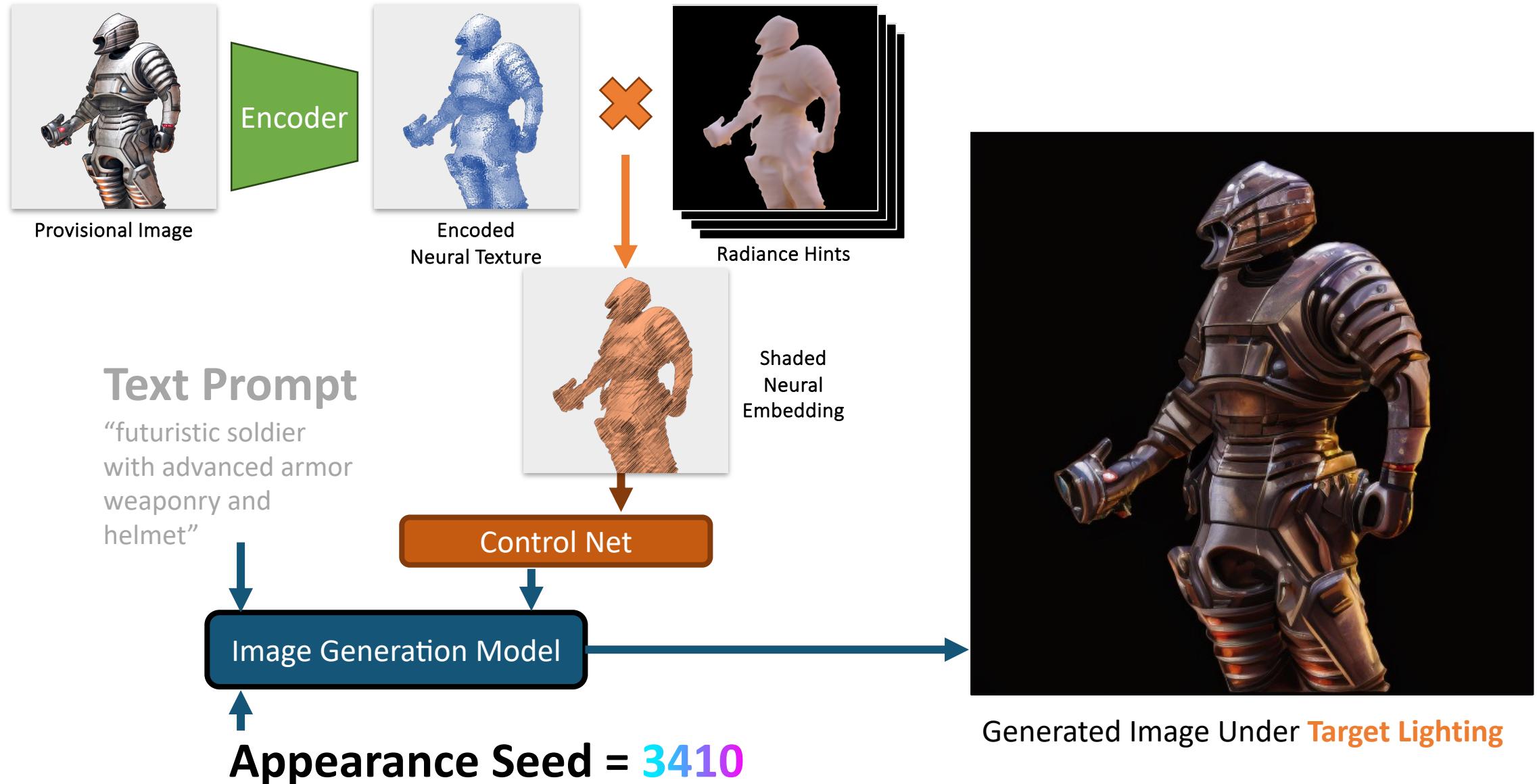


Final Shaded Image

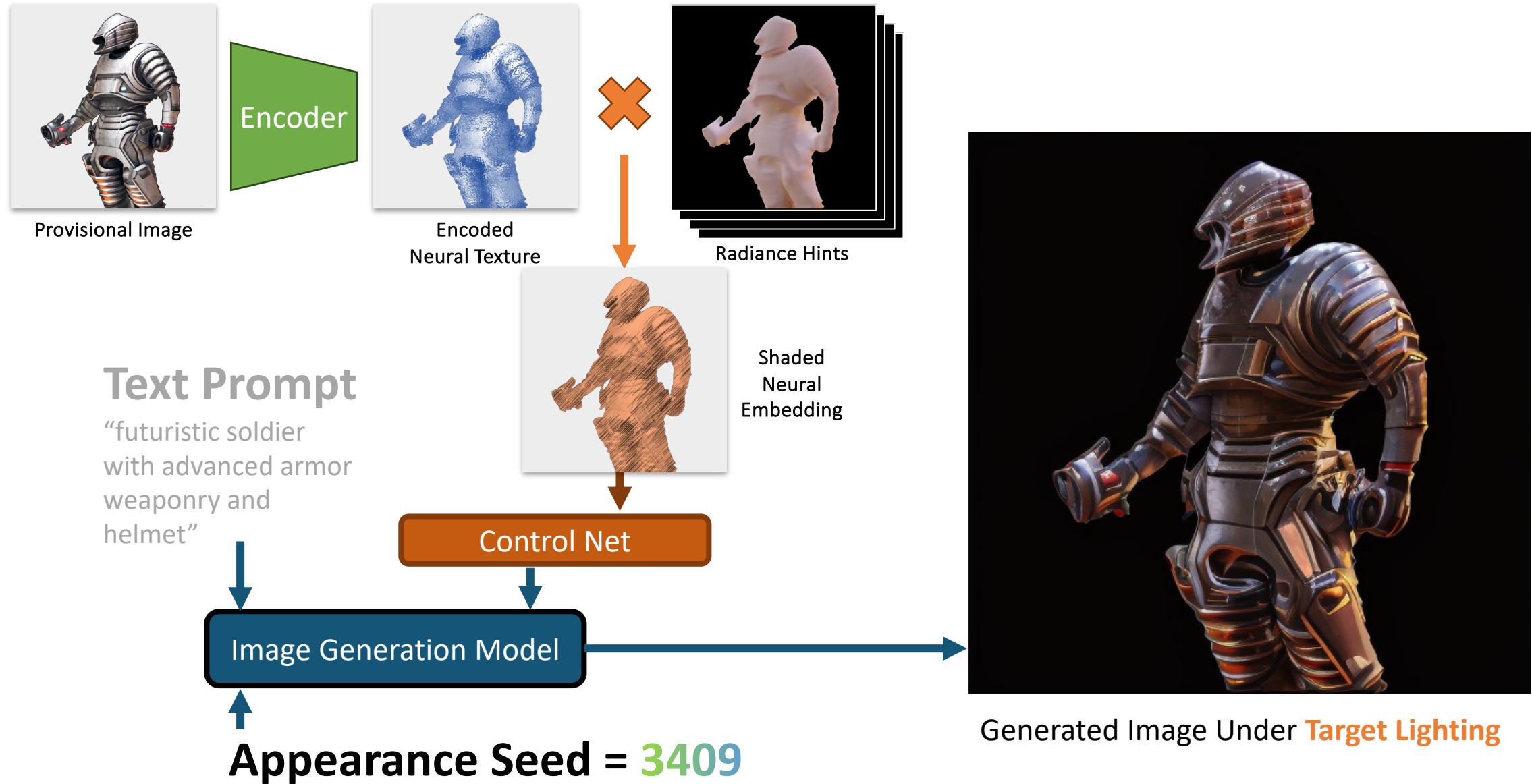
Generate Final Shading Effects via a Trained ControlNet



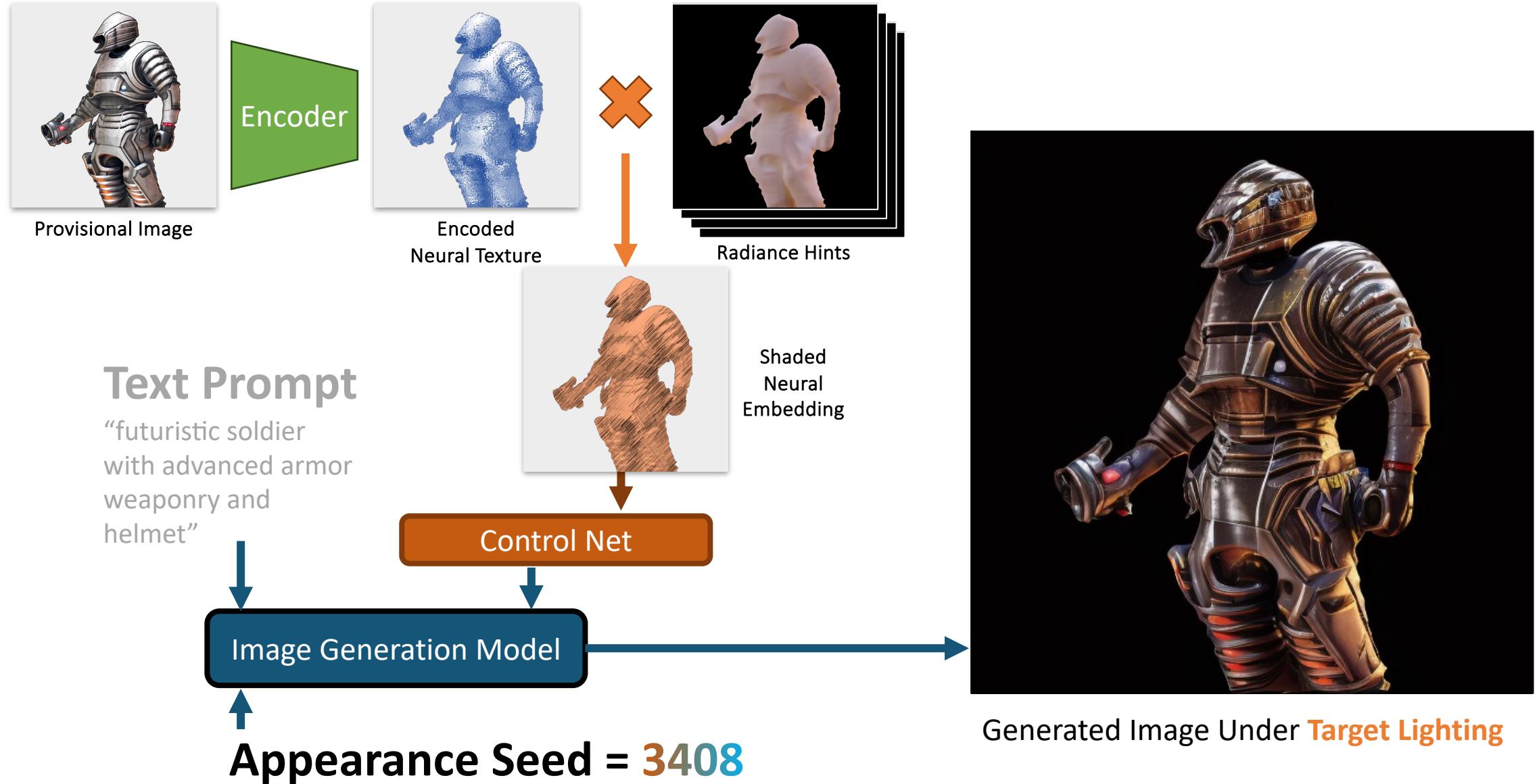
Generate Final Shading Effects via a Trained ControlNet



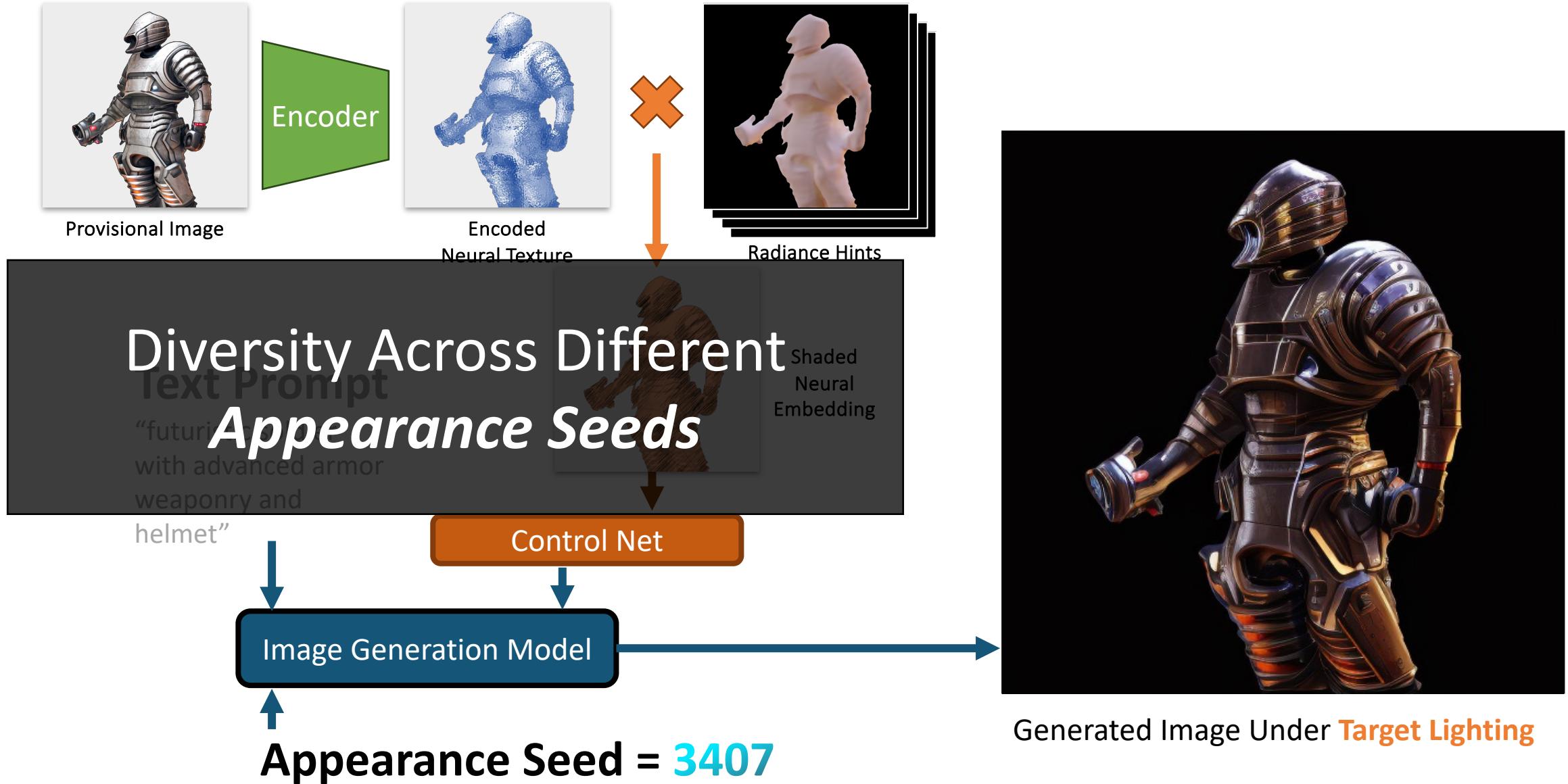
Generate Final Shading Effects via a Trained ControlNet



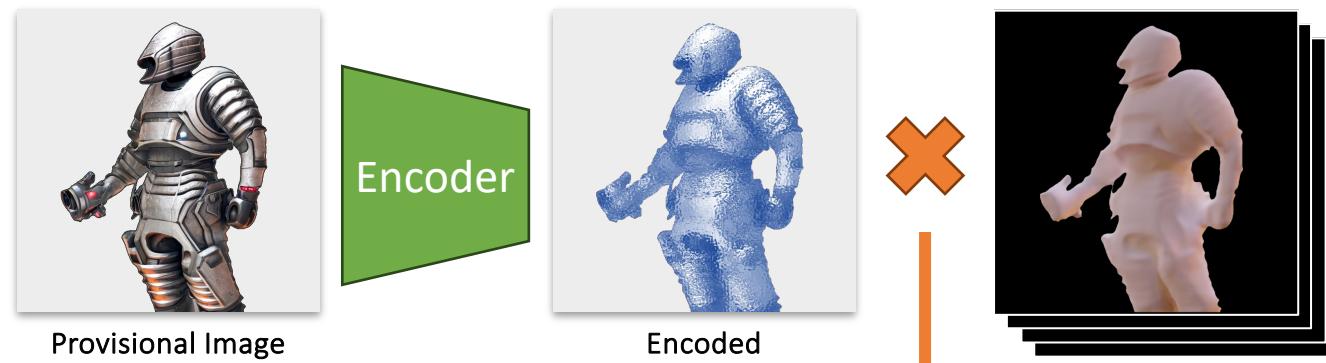
Generate Final Shading Effects via a Trained ControlNet



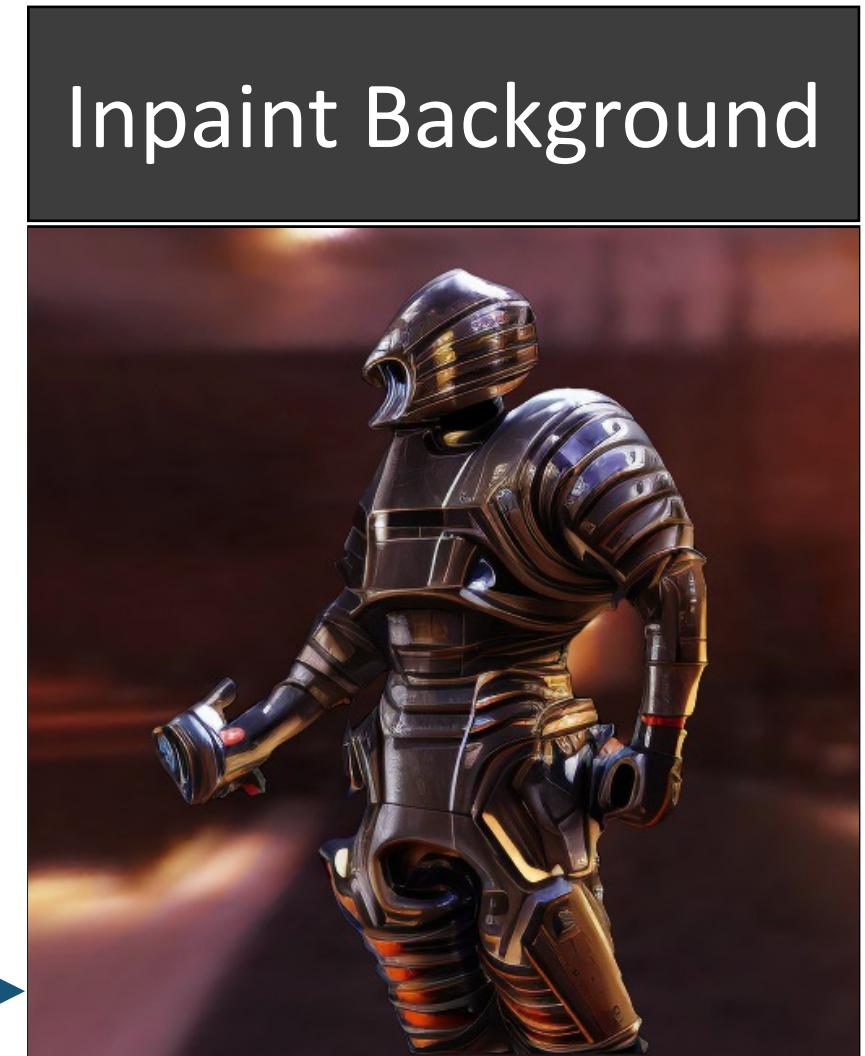
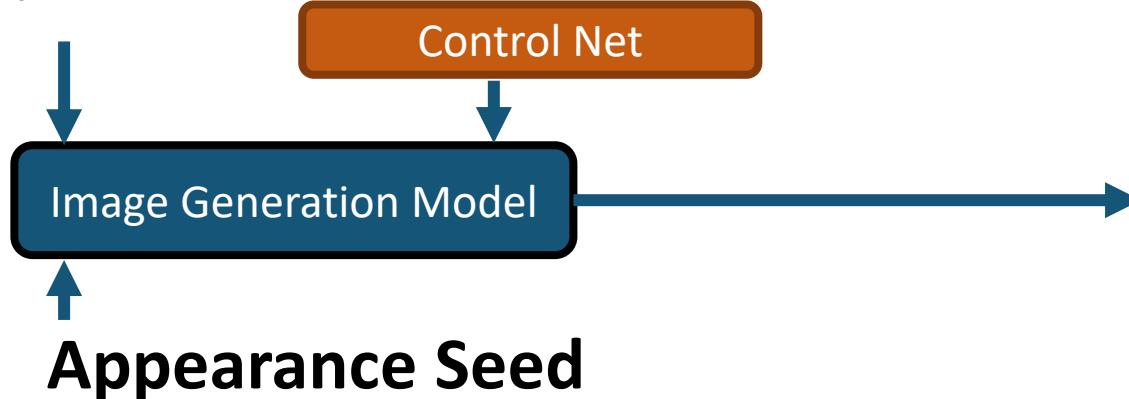
Generate Final Shading Effects via a Trained ControlNet



Generate Final Shading Effects via a Trained ControlNet

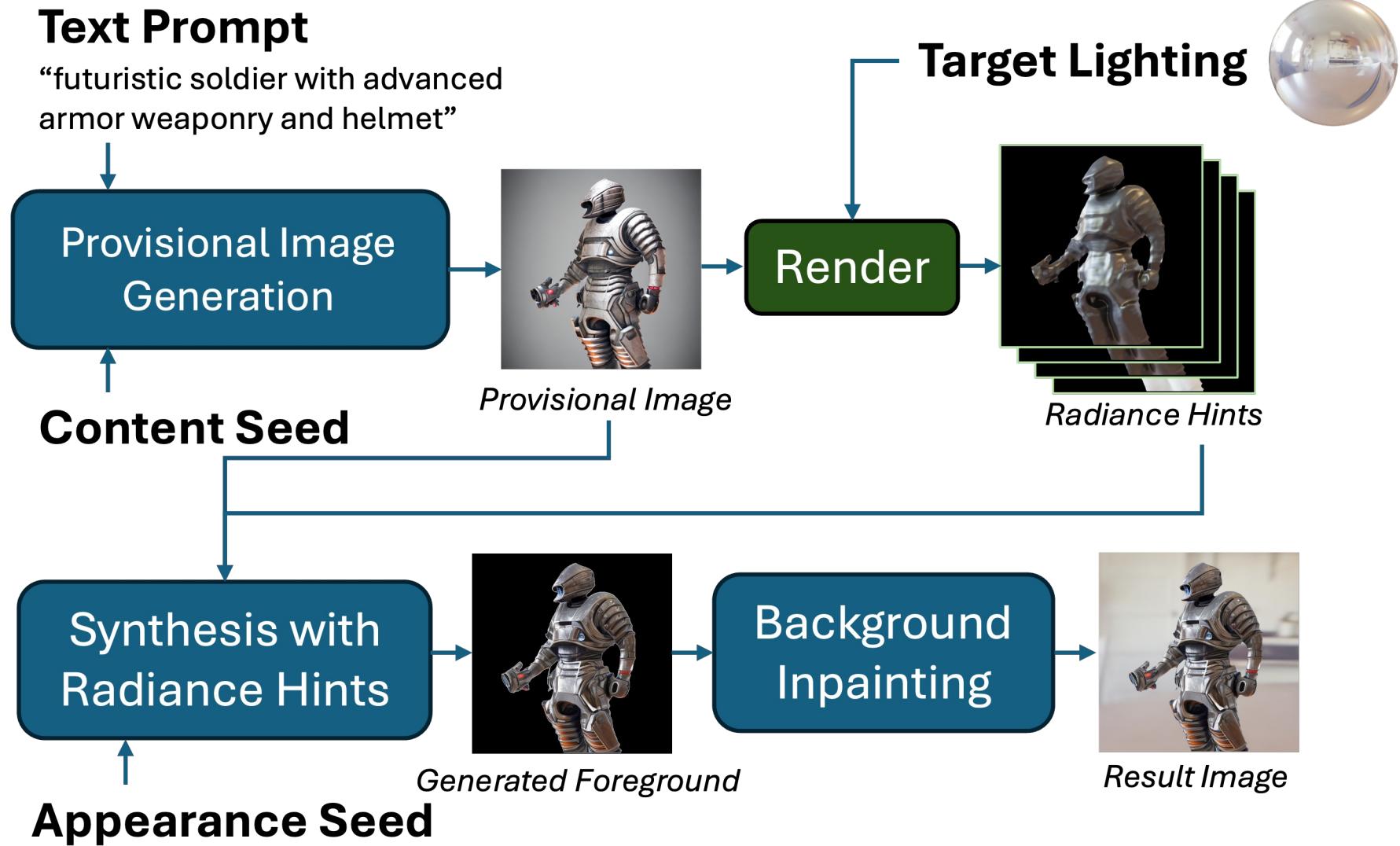


Text Prompt
“futuristic soldier
with advanced armor
weaponry and
helmet”



Generated Image Under **Target Lighting**

Full Pipeline



Training Data – Overview



25K

Shapes with PBR Texture

13K from Objaverse-LVIS
12K material enhanced



4

Views per Shape

12

Lightings per View

Total: 1.2M

Training Image Pairs

Each Contains:

1 Target Image + 4 Radiance Hints

Training Data – Material Enhancement



Random Homogeneous Specular

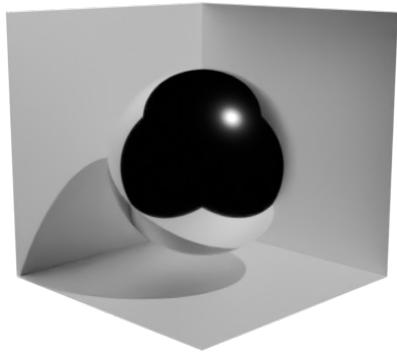


Random Homogeneous Material

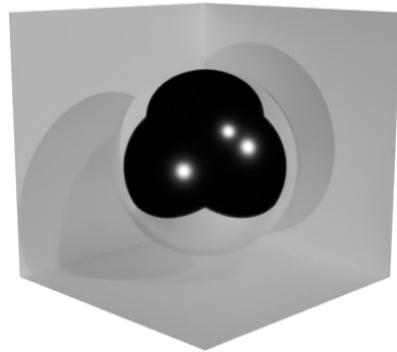


Random SVBRDF

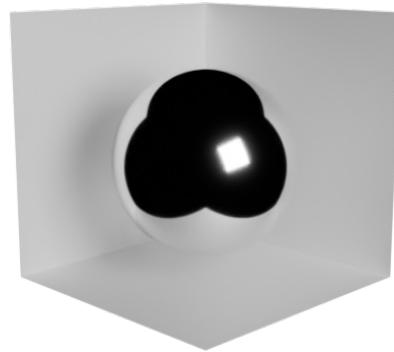
Training Data – Multiple Lightings



Single Point Lighting



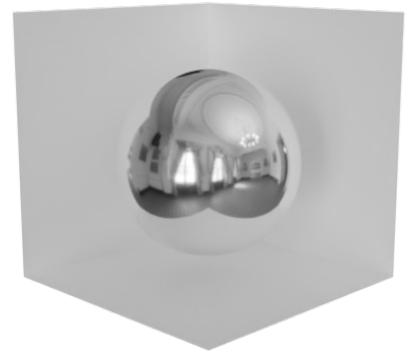
Multiple Point Lightings



Area Lighting



Env Lighting



Monochrome Env Lighting

Results – Basic Lighting Control



“futuristic soldier with advanced armor weaponry and helmet”

“rusty steel toy frog with spatially varying materials with the body diffuse but shiny eyes”

Results – Environmental Lighting



"3D animation character minimal art toy"



"An elephant sculpted from plaster and the elephant nose is decorated with the golden texture"



"a decorated plaster round plate with blue fine silk ribbon around it"

Results – Environmental Lighting



"rusty copper toy frog with spatially varying materials some parts are shinning other parts are rough"



"steampunk space tank with delicate details"



"stone griffin"

Results – Point Lighting



“a large colorful candle, high quality product photo”



“gorgeous ornate fountain made of marble”



“rusty sculpture of a phoenix with its head more polished yet the wings are more rusty”

Results – Diversity Across Appearance Seeds



“leather glove”

Results – Material Control via Prompts



Extension – Depth-conditioned Gen w/ GT Mesh



“a wolf head sculpture and a vase on a desk”

Conclusion

- A novel method for exerting lighting control during text-driven diffusion-based image generation
- **Fine-grained control for arbitrary lighting condition**



Project Page

Future Work

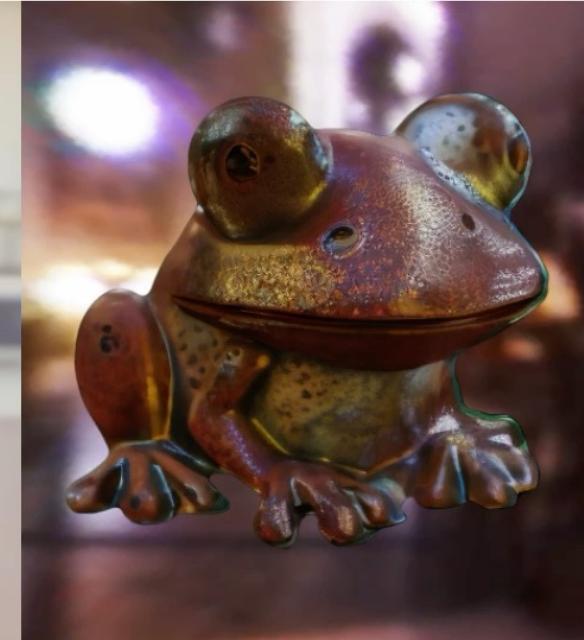
- Real world material estimation with generative prior
- Image-to-3D with material
- Large-scale scanned object PBR dataset

Thanks for Listening!

DiLightNet: Fine-grained Lighting Control for Diffusion-based Image Generation
<https://dilightnet.github.io/>



Project Page



"futuristic soldier with advanced armor weaponry and helmet"

"rusty steel toy frog with spatially varying materials with the body diffuse but shiny eyes"