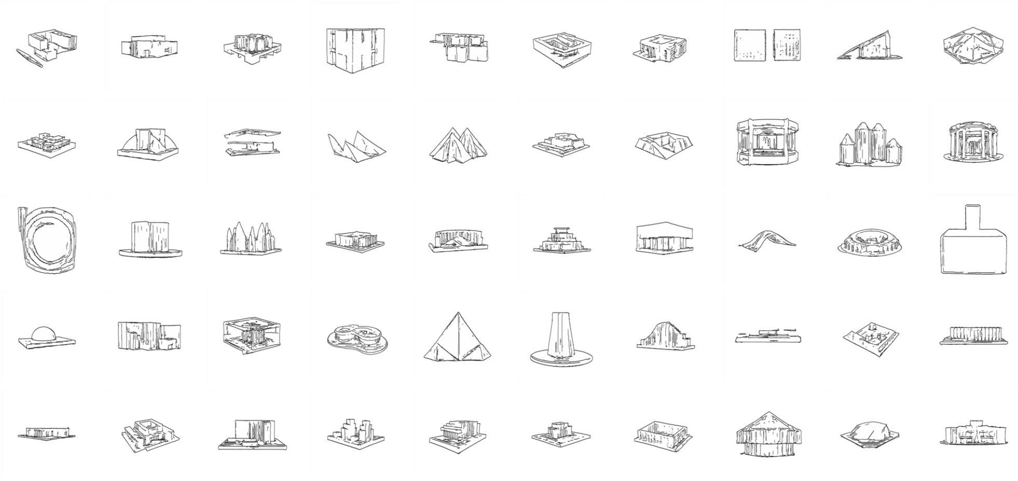
11685- Introduction to Deep learning Project



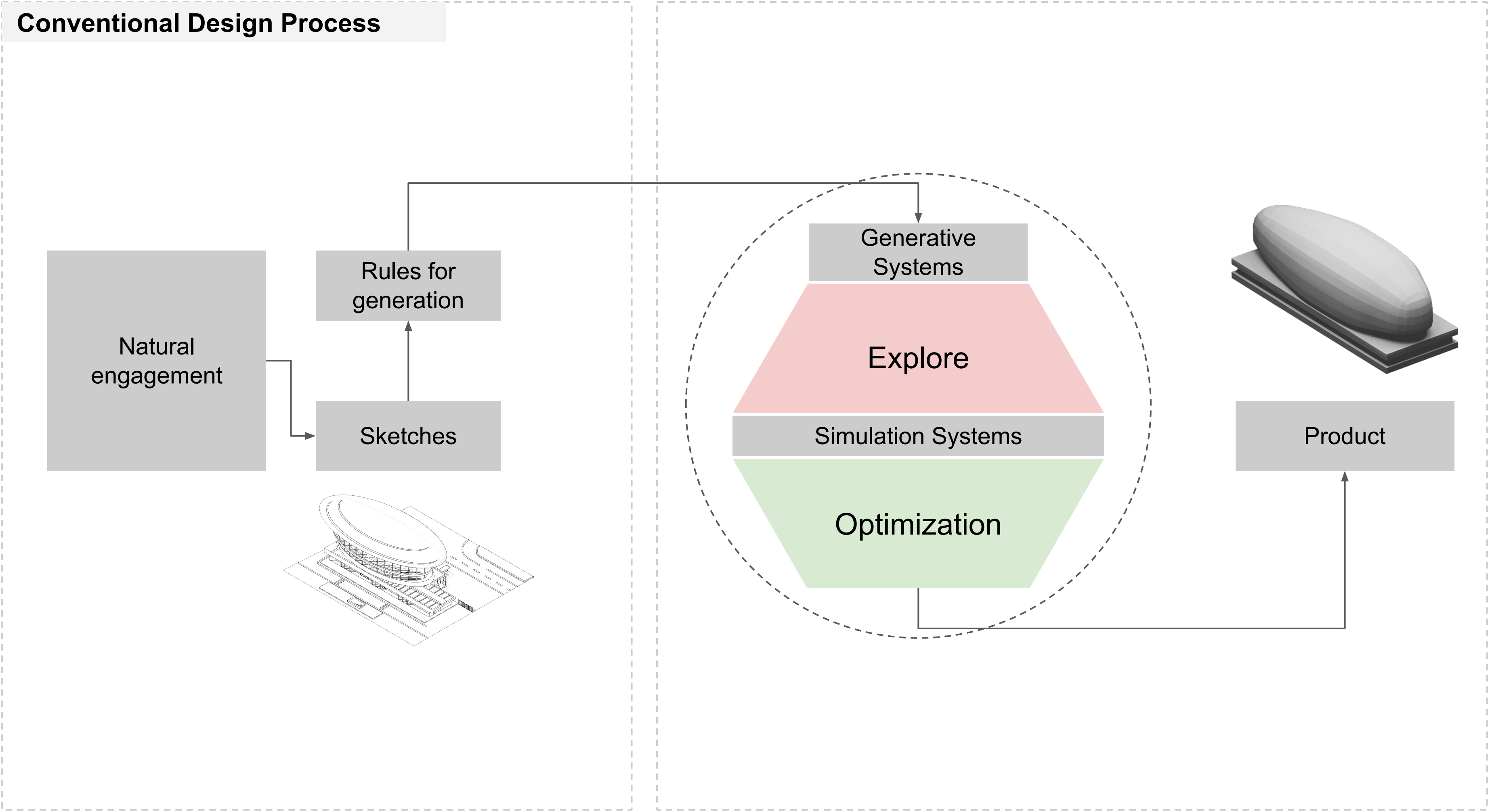
3Ts_Model for Architectural design process

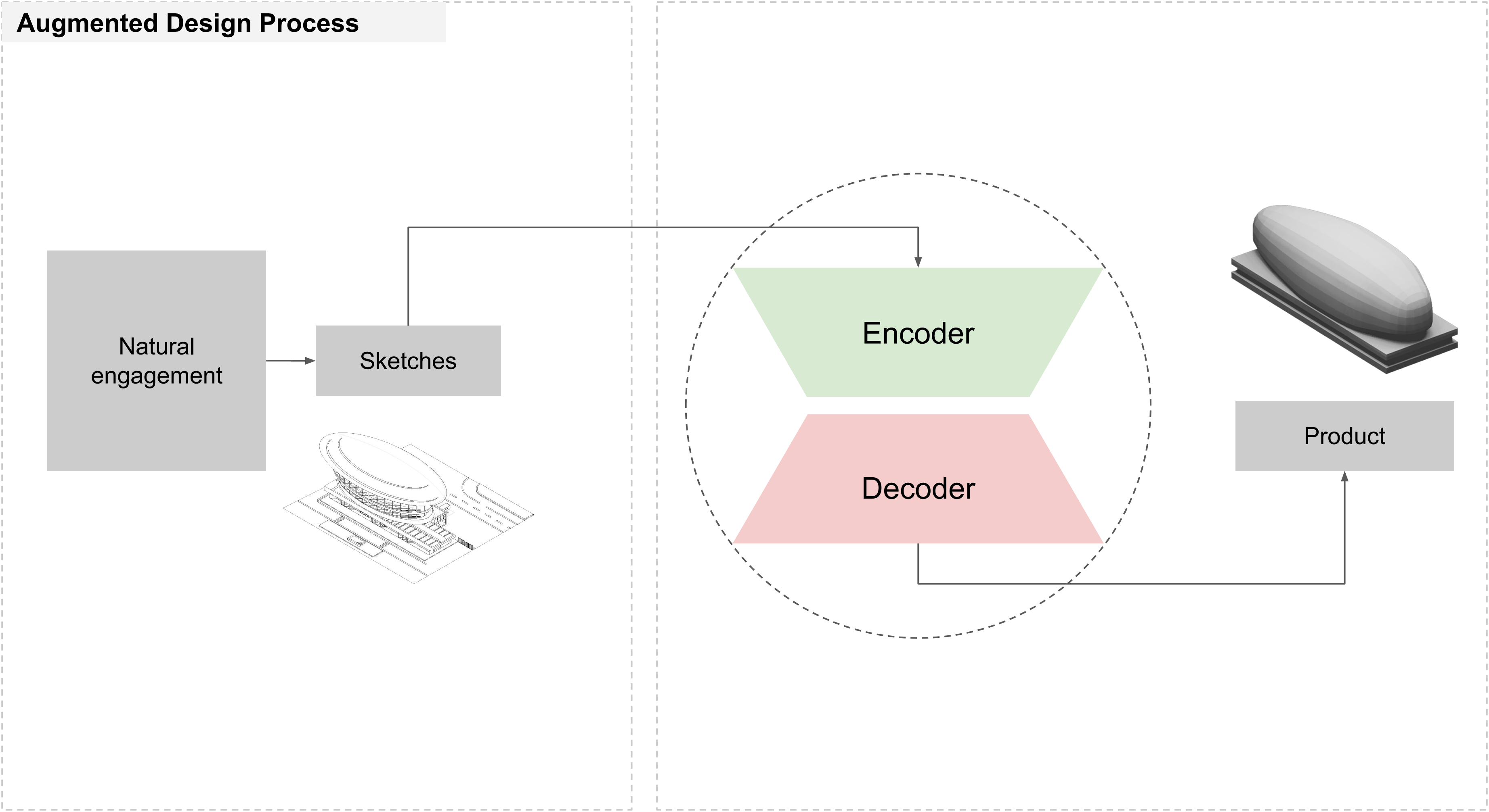
David Chen
School of Architecture
Carnegie Mellon University
Pittsburgh, PA 15213
davidch2@andrew.cmu.edu

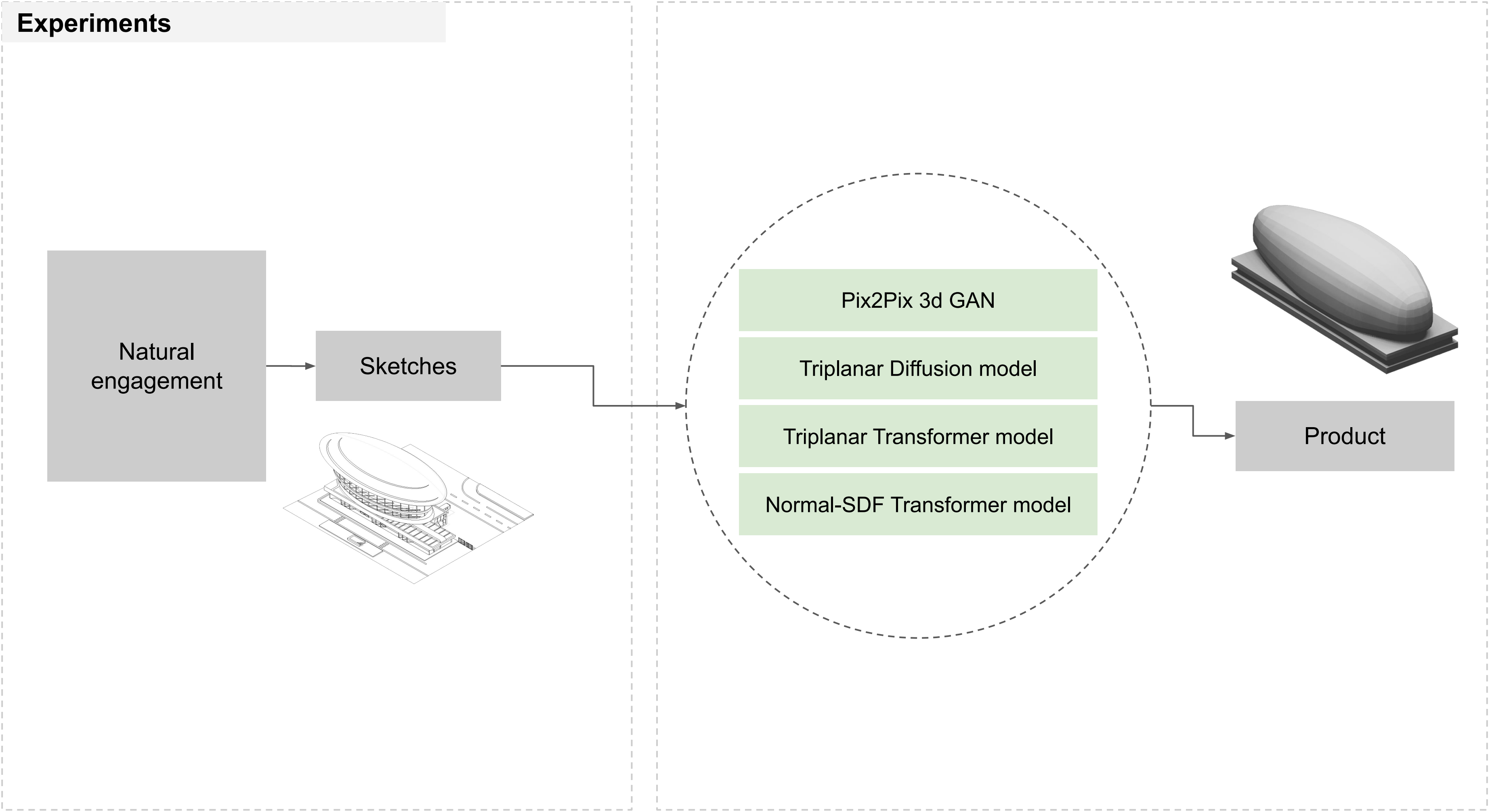
Chia Hui Yen
School of Architecture
Carnegie Mellon University
Pittsburgh, PA 15213
huiyenc@andrew.cmu.edu

Graham Felton
School of Architecture
Carnegie Mellon University
Pittsburgh, PA 15213
gtf@andrew.cmu.edu

Karthick Raja
School of Architecture
Carnegie Mellon University
Pittsburgh, PA 15213
kbangaru@andrew.cmu.edu


Conventional Design Process
Natural engagement
Rules for generation
Sketches
Generative Systems
Explore
Simulation Systems
Optimization
Product


Augmented Design Process
Natural engagement
Sketches
Encoder
Decoder
Product


Experiments
Natural engagement
Sketches
Pix2Pix 3d GAN
Triplanar Diffusion model
Triplanar Transformer model
Normal-SDF Transformer model
Product

Triplanar Representation

3D Mesh Object

↓

Projected Feature Maps

↓

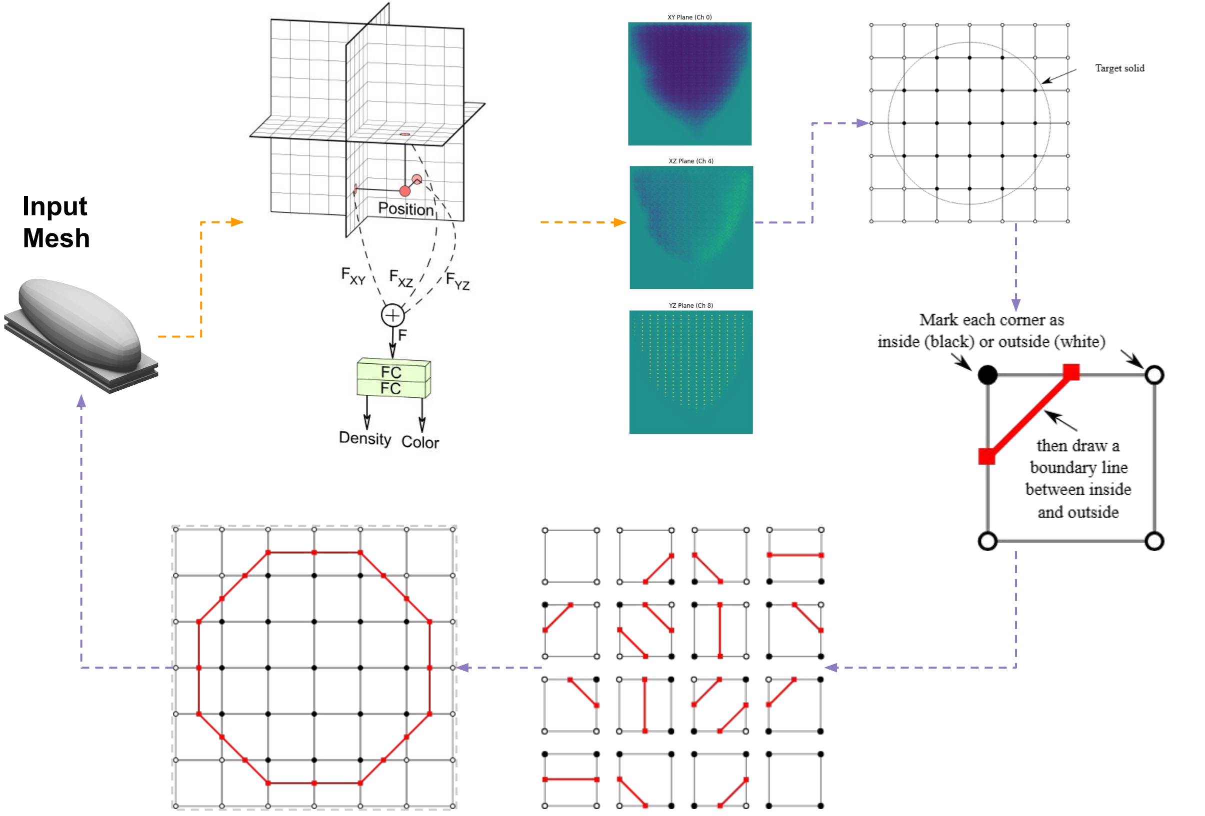
Implicit, Learnable Representation

↓

Decoder

↓

Output 3D Representation

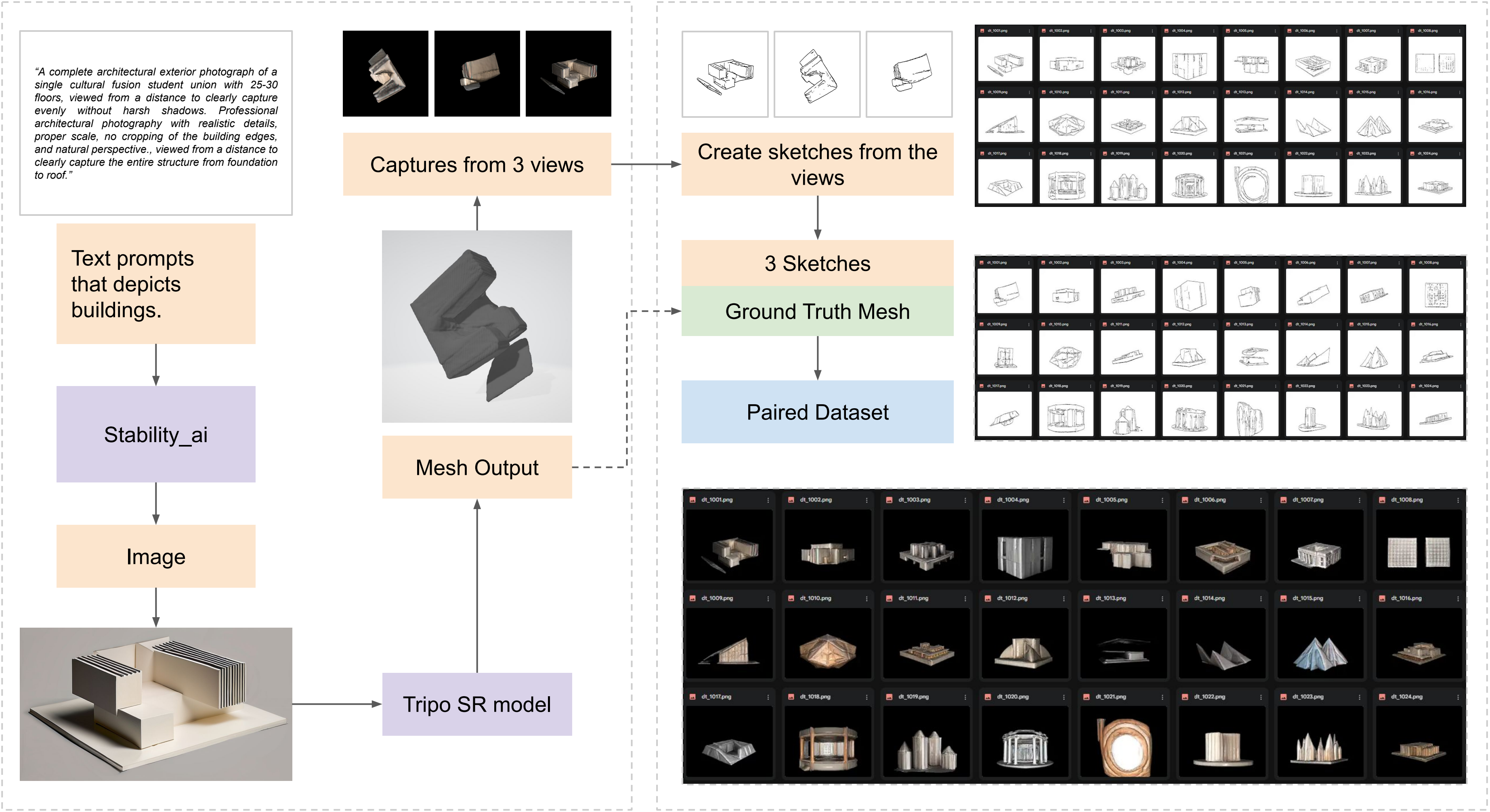


Triplanar representation

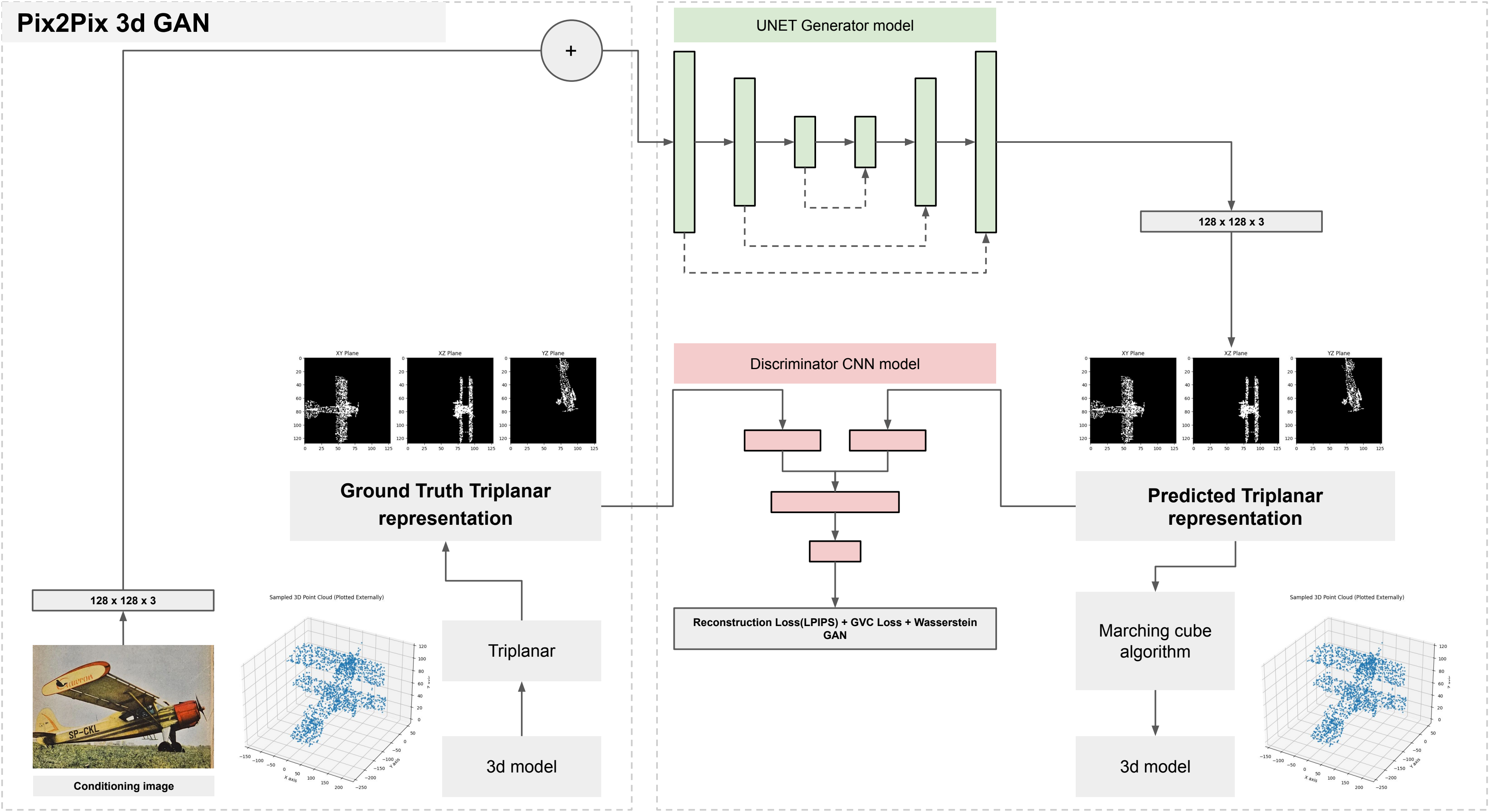
Marching Cube Algorithm

Image from: https://www.boristhebrave.com/2018/04/15/marching-cubes-tutorial/

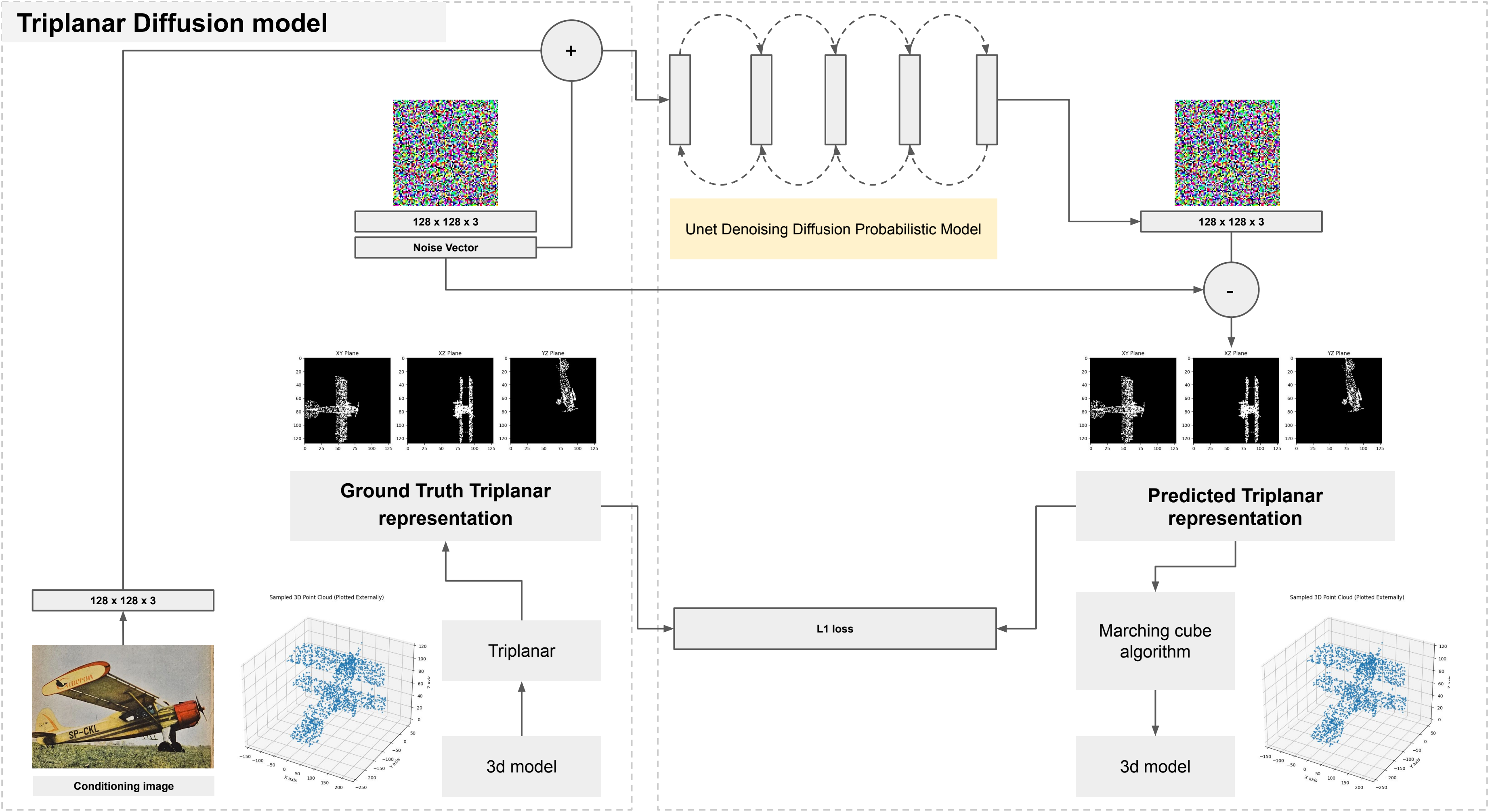
Image from: https://doi.org/10.48550/arXiv.2112.07945



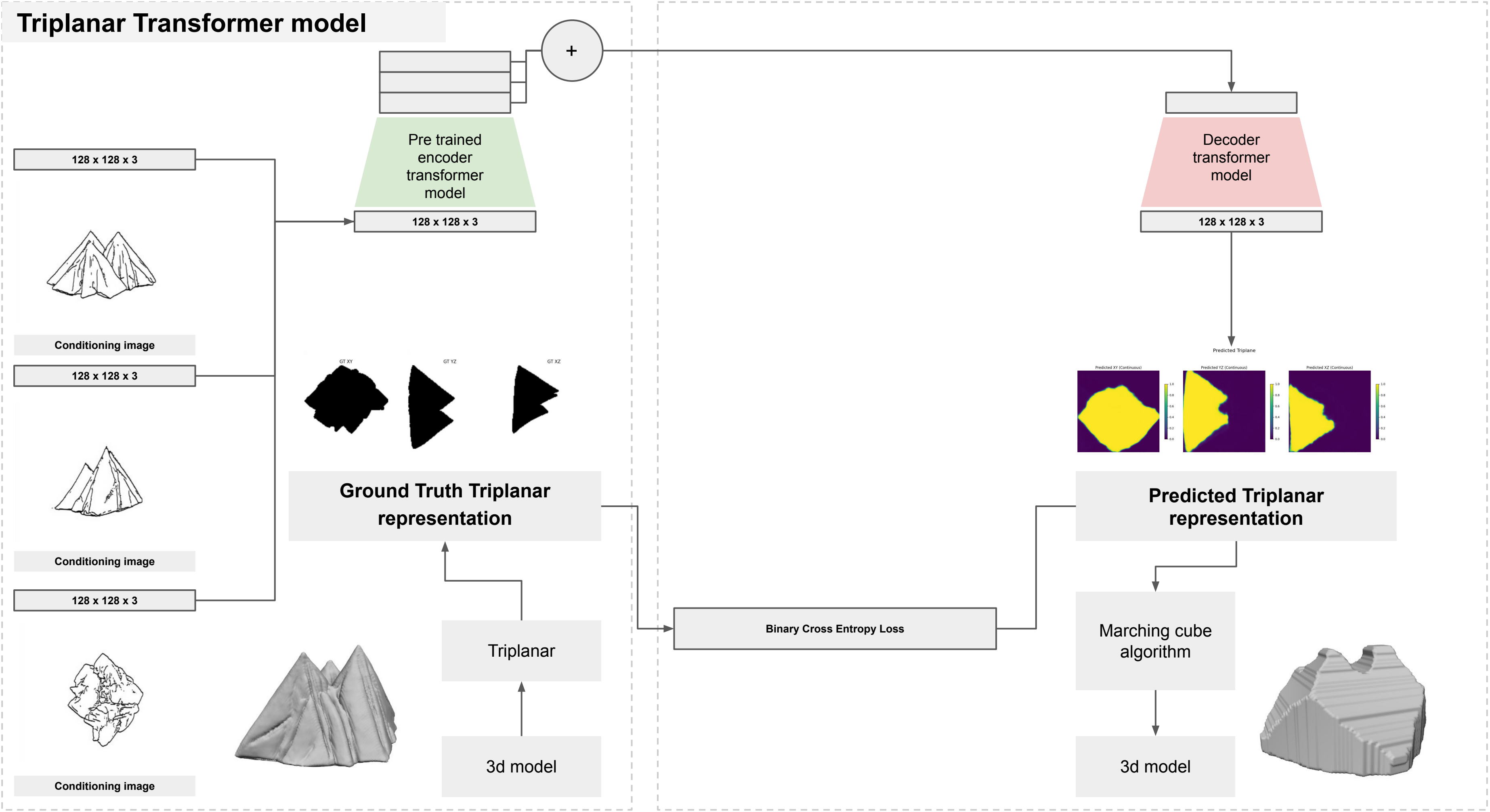
"A complete architectural exterior photograph of a single cultural fusion student union with 25-30 floors, viewed from a distance to clearly capture evenly without harsh shadows. Professional architectural photography with realistic details, proper scale, no cropping of the building edges, and natural perspective., viewed from a distance to clearly capture the entire structure from foundation to roof."
Text prompts that depicts buildings.
Stability_ai
Image
Tripo SR model
Mesh Output
Captures from 3 views
Create sketches from the views
3 Sketches
Ground Truth Mesh
Paired Dataset
dt_1001.png
dt_1002.png
dt_1003.png
dt_1004.png
dt_1005.png
dt_1006.png
dt_1007.png
dt_1008.png
dt_1009.png
dt_1010.png
dt_1011.png
dt_1012.png
dt_1013.png
dt_1014.png
dt_1015.png
dt_1016.png
dt_1017.png
dt_1018.png
dt_1019.png
dt_1020.png
dt_1021.png
dt_1022.png
dt_1023.png
dt_1024.png

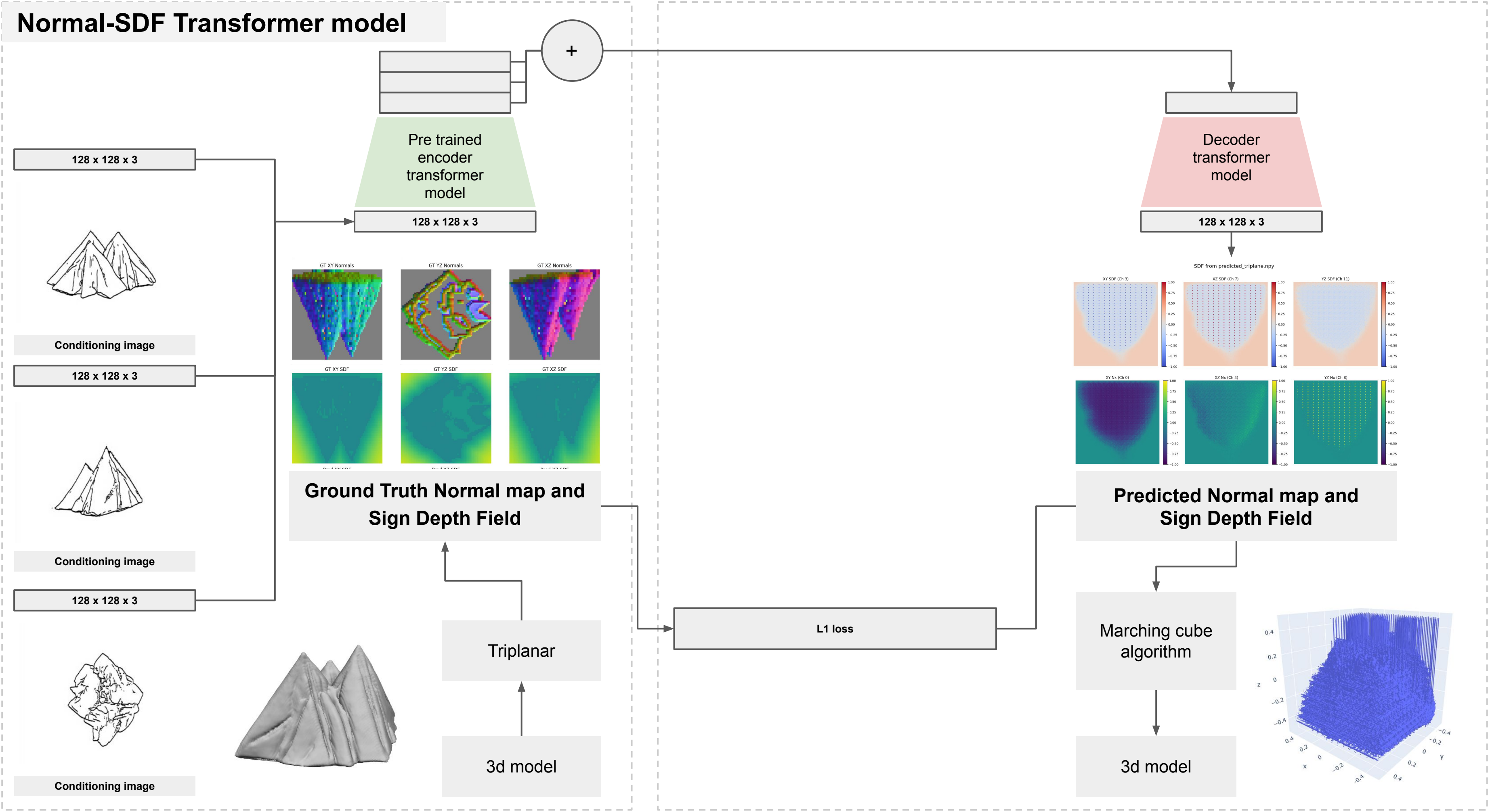


Pix2Pix 3d GAN
+
UNET Generator model
128 x 128 x 3
Discriminator CNN model
XY Plane
XZ Plane
YZ Plane
Ground Truth Triplanar representation
Predicted Triplanar representation
Reconstruction Loss(LPIPS) + GVC Loss + Wasserstein GAN
128 x 128 x 3
Sampled 3D Point Cloud (Plotted Externally)
Triplanar
3d model
Marching cube algorithm
3d model
Conditioning image



Triplanar Diffusion model
+
128 x 128 x 3
Noise Vector
Unet Denoising Diffusion Probabilistic Model
128 x 128 x 3
-
XY Plane
XZ Plane
YZ Plane
Ground Truth Triplanar representation
Predicted Triplanar representation
128 x 128 x 3
Sampled 3D Point Cloud (Plotted Externally)
Triplanar
L1 loss
Marching cube algorithm
Conditioning image
3d model
3d model


Triplanar Transformer model
+
Pre trained encoder transformer model
128 x 128 x 3
128 x 128 x 3
Conditioning image
128 x 128 x 3
Conditioning image
128 x 128 x 3
Conditioning image
Ground Truth Triplanar representation
Triplanar
3d model
Decoder transformer model
128 x 128 x 3
Predicted Triplanar representation
Binary Cross Entropy Loss
Marching cube algorithm
3d model



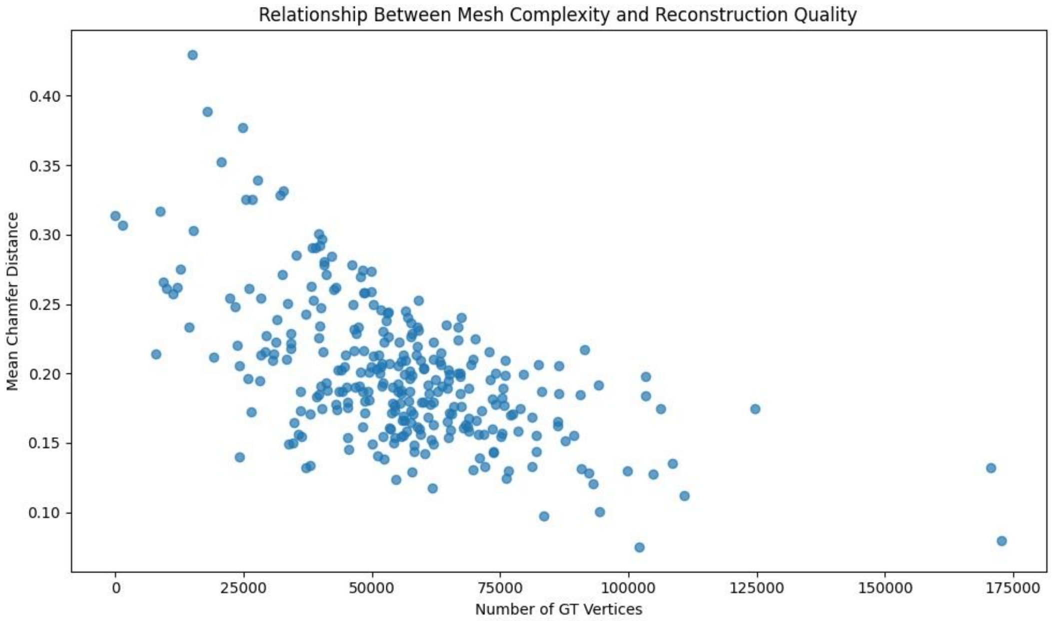
Normal-SDF Transformer model
+
Pre trained encoder transformer model
128 x 128 x 3
128 x 128 x 3
Conditioning image
128 x 128 x 3
Conditioning image
128 x 128 x 3
Conditioning image
GT XY Normals
GT YZ Normals
GT XZ Normals
GT XY SDF
GT YZ SDF
GT XZ SDF
Ground Truth Normal map and Sign Depth Field
Triplanar
3d model
L1 loss
Decoder transformer model
128 x 128 x 3
SDF from predicted_triplane.npy
XY SDF (Ch 3)
XZ SDF (Ch 7)
YZ SDF (Ch 11)
XY Nx (Ch 0)
XZ Nx (Ch 4)
YZ Nx (Ch 8)
Predicted Normal map and Sign Depth Field
Marching cube algorithm
3d model

Chamfer distance Evaluation (Triplanar Transformer model)

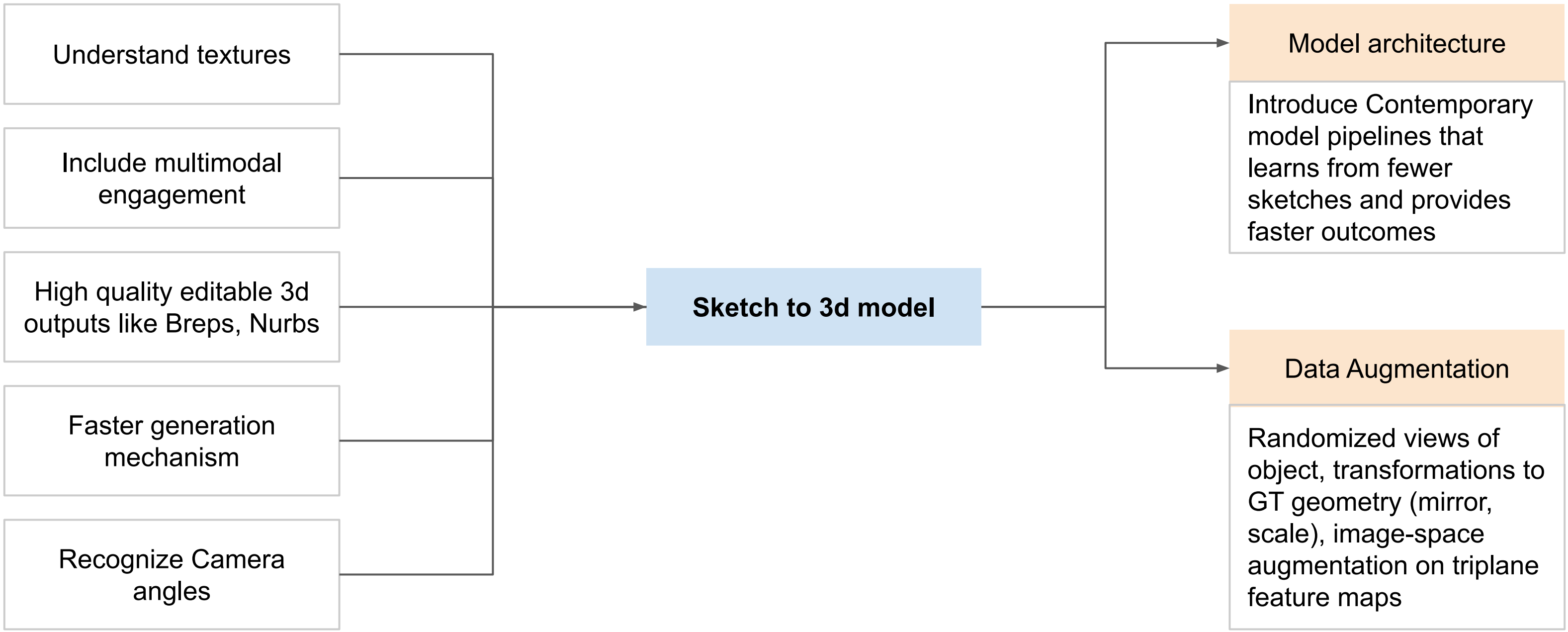
- ***Chamfer Distance*** *is a common metric for evaluating the similarity between two 3D shapes.*
- *We compute it by uniformly sampling 10,000 surface points from both the predicted and ground-truth meshes, then averaging the nearest-neighbor distances in both directions.This bidirectional distance reflects how closely the reconstructed geometry aligns with the original surface.*

Chamfer Distance evaluation on Validation dataset	
Total_models_evaluated	300
Mean_chamfer_dist	0.200
Median_chamfer_dist	0.192
Min_chamfer_dist	0.075
Max_chamfer_dist	0.429

Other Models	**Chamfer Distance**
One-2-3-45	0.227
ZeroShape	0.160
TGS	0.122
OpenLRM	0.180
TripoSR	0.111
Triplanar Transformer model	**0.200**



Future Proposals and measures



11685- Introduction to Deep learning Project

THANK YOU

David Chen
School of Architecture
Carnegie Mellon University
Pittsburgh, PA 15213
davidch2@andrew.cmu.edu

Chia Hui Yen
School of Architecture
Carnegie Mellon University
Pittsburgh, PA 15213
huiyenc@andrew.cmu.edu

Graham Felton
School of Architecture
Carnegie Mellon University
Pittsburgh, PA 15213
gtf@andrew.cmu.edu

Karthick Raja
School of Architecture
Carnegie Mellon University
Pittsburgh, PA 15213
kbangaru@andrew.cmu.edu

References

1. Chan, C., Durand, F., & Isola, P. (2022). Learning to generate line drawings that convey geometry and semantics.
2. Deng, K., Yang, G., Ramanan, D., & Zhu, J.-Y. (2023). 3D-aware conditional image synthesis.
3. Mao, A., Mohri, M., & Zhong, Y. (2023). Cross-entropy loss functions: Theoretical analysis and applications.
4. Oquab, M., Darcet, T., Moutakanni, T., Vo, H., Szafraniec, M., Khalidov, V., et al. (2024). Dinov2: Learning robust visual features without supervision.
5. Park, J. J., Florence, P., Straub, J., Newcombe, R., & Lovegrove, S. (2019). DeepSDF: Learning continuous signed distance functions for shape representation.
6. Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., et al. (2017). Automatic differentiation in PyTorch.
7. Podell, D., English, Z., Lacey, K., Blattmann, A., Dockhorn, T., Müller, J., Penna, J., & Rombach, R. (2023). SDXL: Improving latent diffusion models for high-resolution image synthesis
8. Shue, J. R., Chan, E. R., Po, R., Ankner, Z., Wu, J., & Wetzstein, G. (2022). 3D neural field generation using triplane diffusion.
9. Spiess, F., Waltenspül, R., & Schuldt, H. (2024). The Sketchfab 3D Creative Commons Collection (S3D3C).
10. Tochilkin, D., Pankratz, D., Liu, Z., Huang, Z., Letts, A., Li, Y., et al. (2024). TripoSR: Fast 3D object reconstruction from a single image.
11. Dmitry Tochilkin, David Pankratz, Zexiang Liu, Zixuan Huang, Adam Letts, Yangguang Li, Ding Liang, Christian Laforte, Varun Jampani, and Yan-Pei Cao. Triposr: Fast 3d object reconstruction from a single image, 2024.
12. Wu, Z., Song, S., Khosla, A., Yu, F., Zhang, L., Tang, X., & Xiao, J. (2015). 3D ShapeNets: A deep representation for volumetric shapes.
13. Zhuang, X., Huang, Z., Zeng, W., & Caldas, L. (2024). Encoding urban ecologies: Automated building archetype generation through self-supervised learning for energy modeling.
14. Zhuang, X., Ju, Y., Yang, A., & Caldas, L. (2023). Synthesis and generation for 3D architecture volume with generative modeling. *International Journal of Architectural Computing*, *21*(2), 297–314.