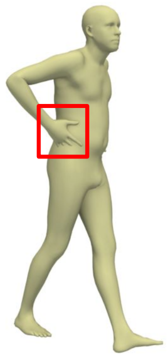
N-Penetrate: Active Learning of Neural Collision Handler for Complex 3D Mesh Deformations

Qingyang Tan[1] Zherong Pan[2] Breannan Smith[3] Takaaki Shiratori[3] Dinesh Manocha[1]
[1] Department of Computer Science, University of Maryland at College Park
[2] Lightspeed & Quantum Studio, Tencent America
[3] Meta Reality Labs Research

Learned 3D Representation with Collision-Free Constraints

Using learned 3D mesh representations to generate new models

Backbone for many motion capture and shape analysis tasks



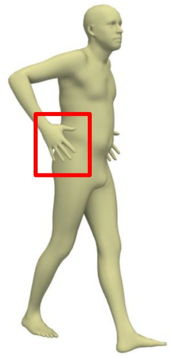
Motivation: Correct 3D mesh models should be collision-free!

Goal: Eliminating self-collisions generated by learning models

Learned 3D Representation with Collision-Free Constraints

Using learned 3D mesh representations to generate new models

Backbone for many motion capture and shape analysis tasks



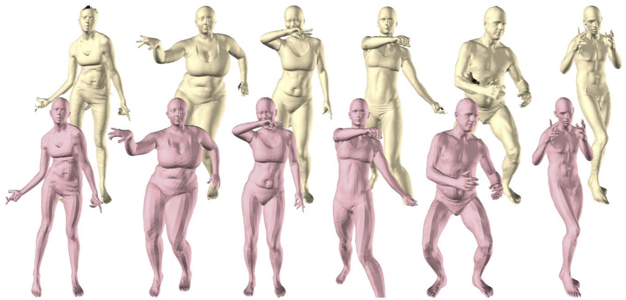
Motivation: Correct 3D mesh models should be collision-free!

Goal: Eliminating self-collisions generated by learning models

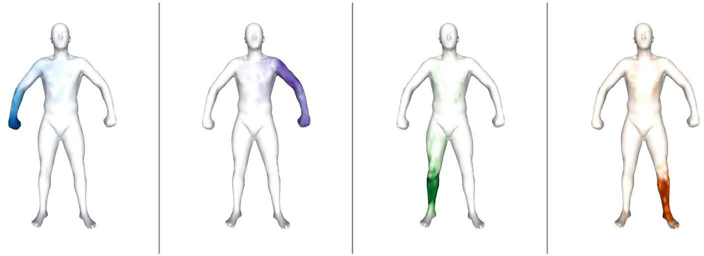
Prior Work & Limitations

Learned 3D model embedding

SMPL (Loper et al. 2015): Human body pose



Neural embedding (Yang et al. 2020): Arbitrary 3D object



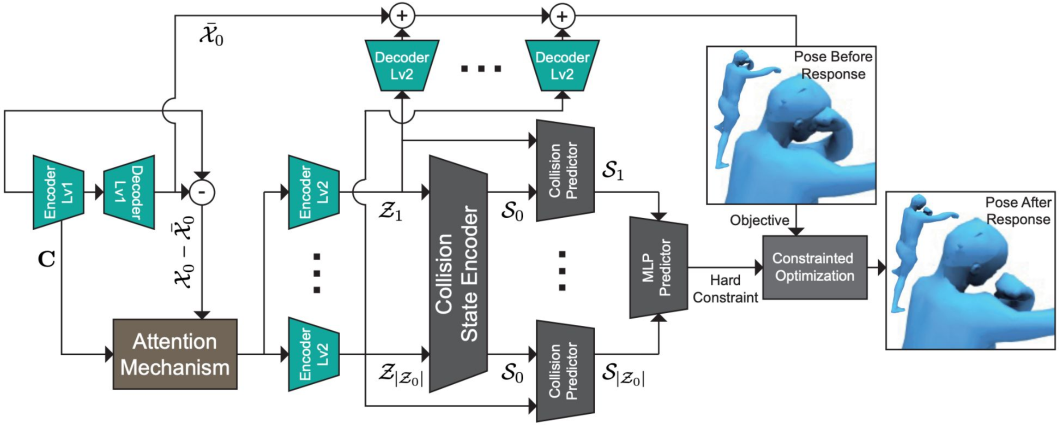
Our method works with all of them!

Prior Work & Limitations

Neural Collision Detector (Tan et al. 2021)

Pro: First neural collision handler for general mesh embedding system

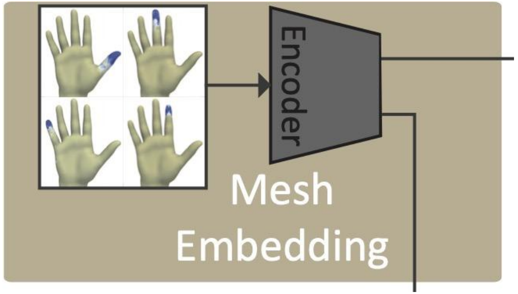
Con: Limited accuracy of collision handler using supervised learning



Proposed Method

Active learning system of collision handler

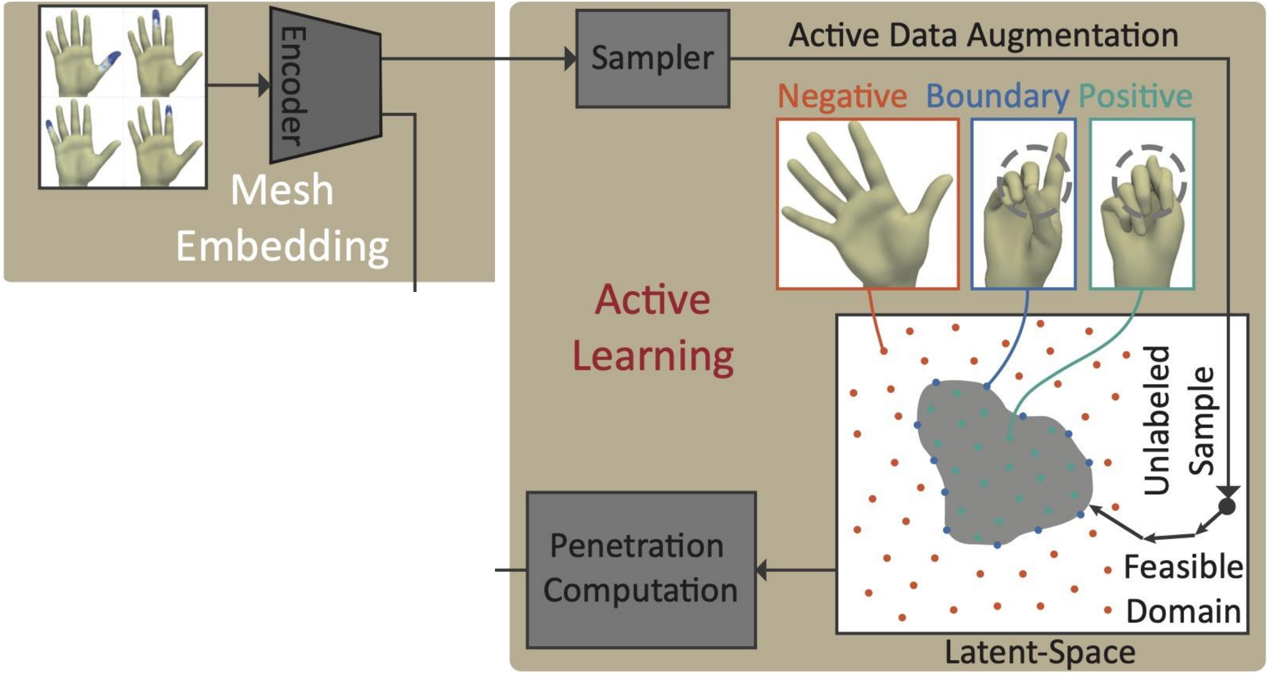
Arbitrary mesh embedding



Proposed Method

Active learning system of collision handler

Arbitrary mesh embedding



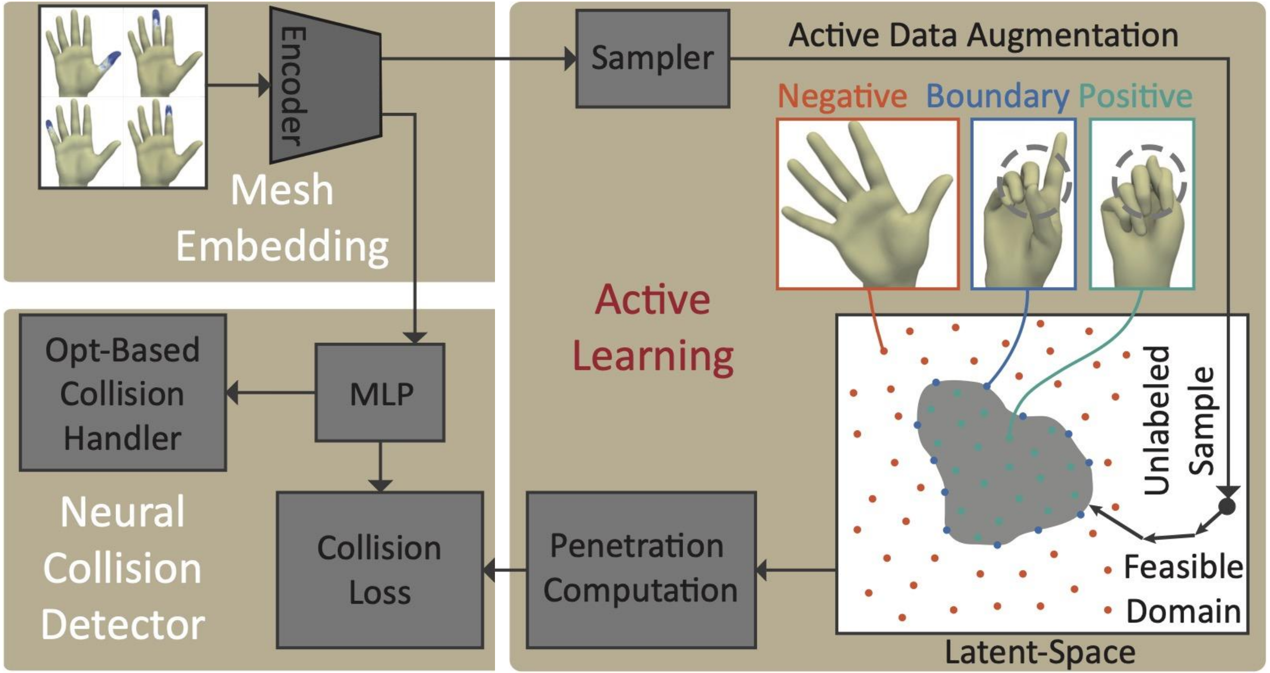
Actively sample on and around decision boundary

Proposed Method

Active learning system of collision handler

Arbitrary mesh embedding

Special loss function for different sample types



Actively sample on and around decision boundary

Active Sampling on the Decision Boundary

Project samples to the 0.5-level-set (the decision boundary)

Solve the unconstrained optimization using quasi-Newton method

$$\underset{Z_{\text{all}}}{\operatorname{argmin}} \frac{1}{2} \|\text{MLP}_c(Z_{\text{all}}, \theta_C) - 0.5\|^2$$

$$Z_{\text{all}} = Z_{\text{all}} - \bar{H}^{-1} \nabla \text{MLP}_c(\text{MLP}_c - 0.5)$$

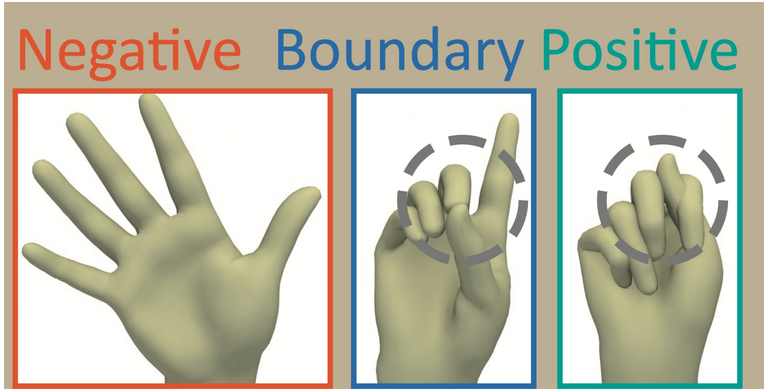
Different Loss Terms for Different Subsets

Use the L_1-loss function for the boundary data to better approximate the decision boundary

$$\mathcal{L}_b = E_{\{Z_{\text{all}} \in \mathcal{D}_b\}}\left[|| \text{MLP}_c(Z_{\text{all}}, \theta_C) - 0.5 || \right]$$

Use the cross-entropy loss for the remaining data

$$\mathcal{L}_{\text{ce}} = E_{\{Z_{\text{all}} \in \mathcal{D}_p \bigcup \mathcal{D}_n\}}\left[\text{CE}(I_c(Z_{\text{all}}), I_c^*(Z_{\text{all}})) \right]$$



Collision Response - Solving Constrained Optimization

Solving the constrained optimization problem using learned neural collision constraint

$$\underset{Z_{\text{all}}}{\text{argmin}}\ \mathbf{E}(Z_{\text{all}}) \quad \text{s.t. } \text{MLP}_c(Z_{\text{all}}, \theta_C) \leq 0.5$$

Given a desired pose, solving for a nearby new pose that is also collision-free using the following function

$$\mathbf{E}(Z_{\text{all}}) = \|Z_{\text{all}} - Z_{\text{all}}^{\text{user}}\|^2/2$$

Evaluation

Datasets

Human pose datasets: SCAPE (Anguelov et al., 2005), MIT Swing & Jump (Vlasic et al., 2008)
Hand dataset from Facebook
Skirt dataset (Yang et al., 2020)
SMPL model: AMASS-MPIMosh (Mahmood et al., 2019)

Baselines

Pure supervised learning method ***Supv*** (Tan et al. 2021)
Adding boundary loss ***Supv+bd***
Adding boundary loss and active learning ***Active+bd***

Evaluation - Accuracy of Collision Detector

Improved collision detection accuracy with the same amount of data

Achieve a **similar accuracy** using a 34.6% **smaller dataset** than ***Supv+bd***

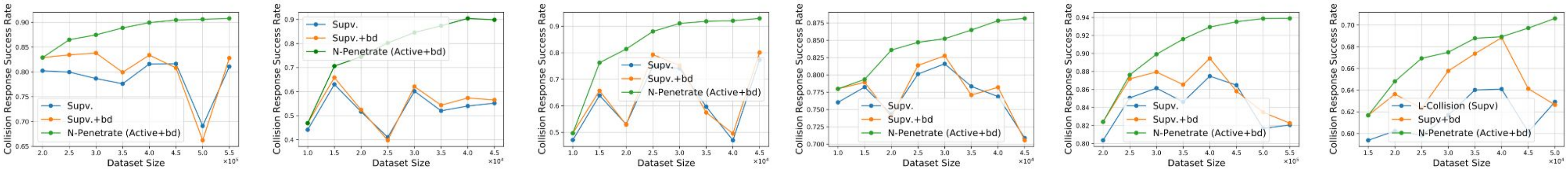
metric	SCAPE	Swing	Jump	Skirt	Hand	AMASS
final dataset size	5.5×10^5	4.5×10^4	4.5×10^4	4×10^4	5.5×10^5	5×10^4
accuracy (Ours (***Active+bd***))	0.9383	0.9638	0.9552	0.9817	0.9692	0.8563
accuracy (***Supv+bd***)	0.9282	0.9609	0.9500	0.9795	0.9650	0.8551
accuracy (***Supv***)	0.9181	0.9460	0.9347	0.9660	0.9558	0.8526
false neg. (Ours (***Active+bd***))	0.05151	0.01485	0.02573	0.01808	0.03582	0.17656
false neg. (***Supv+bd***)	0.05576	0.01766	0.02644	0.01956	0.03652	0.18422
false neg. (***Supv***)	0.06914	0.01969	0.02713	0.02056	0.03727	0.19654
equi. dataset size (***Supv+bd***)	6.63×10^5	7.64×10^4	7.02×10^4	7.71×10^4	7.30×10^5	7.40×10^4

Evaluation - Collision Response Success Rate

Our method has **22.73% more collisions resolved** than Supv+bd

Stable training performance with monotonically increasing accuracy as dataset grows

Supv & ***Supv+bd*** exhibits unstable performance



SCAPE | Swing | Jump | Skirt | Hand | AMASS

Evaluation - Collision Response Quant. Analysis

Our method

mean relative PD reduction - 94.81%

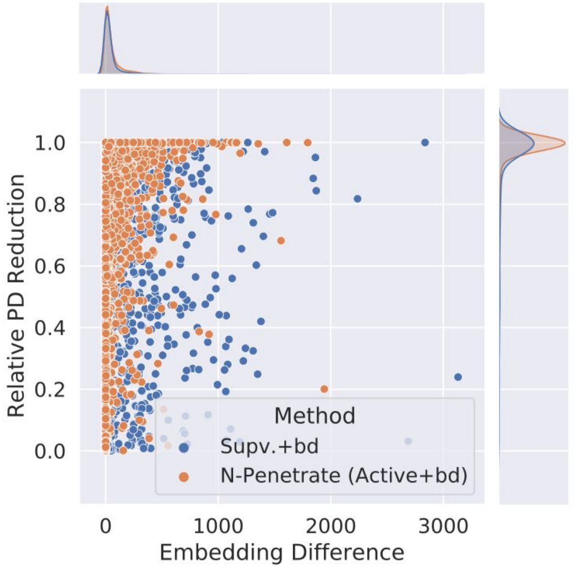
mean embedding difference - 56.17

Supv+bd

mean relative PD reduction - 88.76%

mean embedding difference - 66.11

Our method resolves more collisions while the outputs stay closer to the input latent codes



Thank you!

Qingyang Tan (qytan@umd.edu)
Zherong Pan (zherong.pan.usa@gmail.com)
Breannan Smith (breannan@fb.com)
Takaaki Shiratori (tshiratori@fb.com)
Dinesh Manocha (dmanocha@umd.edu)