A Study on Different Architectures on a 3D Garment Reconstruction Network

by:

Jason Klimack

Supervisors:

Meysam Madadi

Sergio Escalera

Outline

- Introduction
- Overview of CLOTH3D Architecture
- Network Changes
- Experiments
- Conclusions

Introduction

- Applications
  - CGI Characters
  - Fashion
- Realistic garments
  - Proper fitting
  - Cloth dynamics
- 3D Garment reconstruction
  - Create realistic 3D garments

Introduction

- Challenge:
  - Lack of 3D garment data
- 3 Methods to obtain data:
  - 2D Image to 3D model
  - 3D Scans
  - Synthetic Generation
- CLOTH3D Dataset
  - Large-scale synthetic dataset of 3D clothed human sequences
  - 2M+ samples
  - Large variety of garment types, topologies, shapes, sizes, and fabrics
  - SMPL Human body sequences

Introduction

- CLOTH3D Network
  - GCVAE
  - Offsets
  - Skirt/Dress Topology
- Goal:
  - Achieve better understanding of GNNs for clothes simulation
  - Study different architectures for 3D garment reconstruction
- Network Modifications:
  - Pooling
  - Central Difference Convolution
  - Octave Convolution

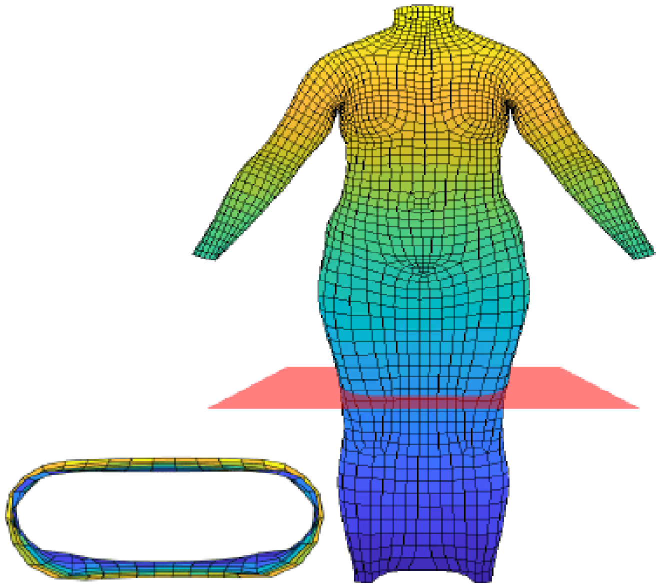
Overview of CLOTH3D Architecture

- CLOTH3D dataset + DL Network
- Challenge:
  - Large variability in Garment type and topology
- Solution:
  - Encode garments as offsets
- Graph Conditional Variational Auto-Encoder (GCVAE)

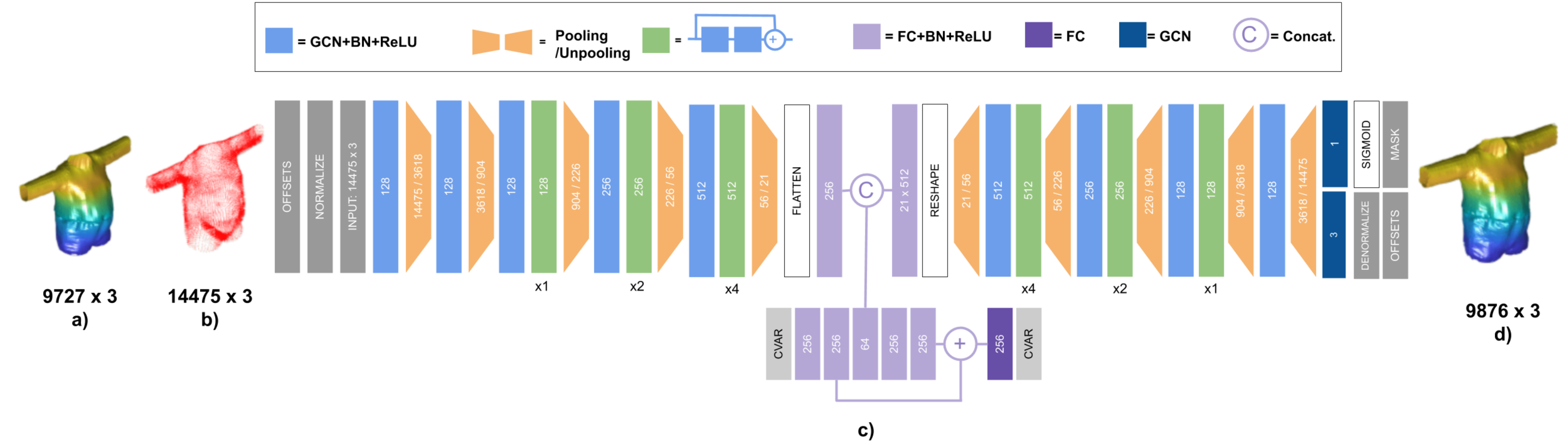
Overview of CLOTH3D Architecture

Pre-Processing

- Represent garments as offsets from SMPL body
- Mask used to highlight garment vertices
- Body-garment matching: non-rigid ICP
- Body topology for skirt/dress



Overview of CLOTH3D Architecture



Overview of CLOTH3D Architecture

Graph Convolution

- Spectral Graph Convolution

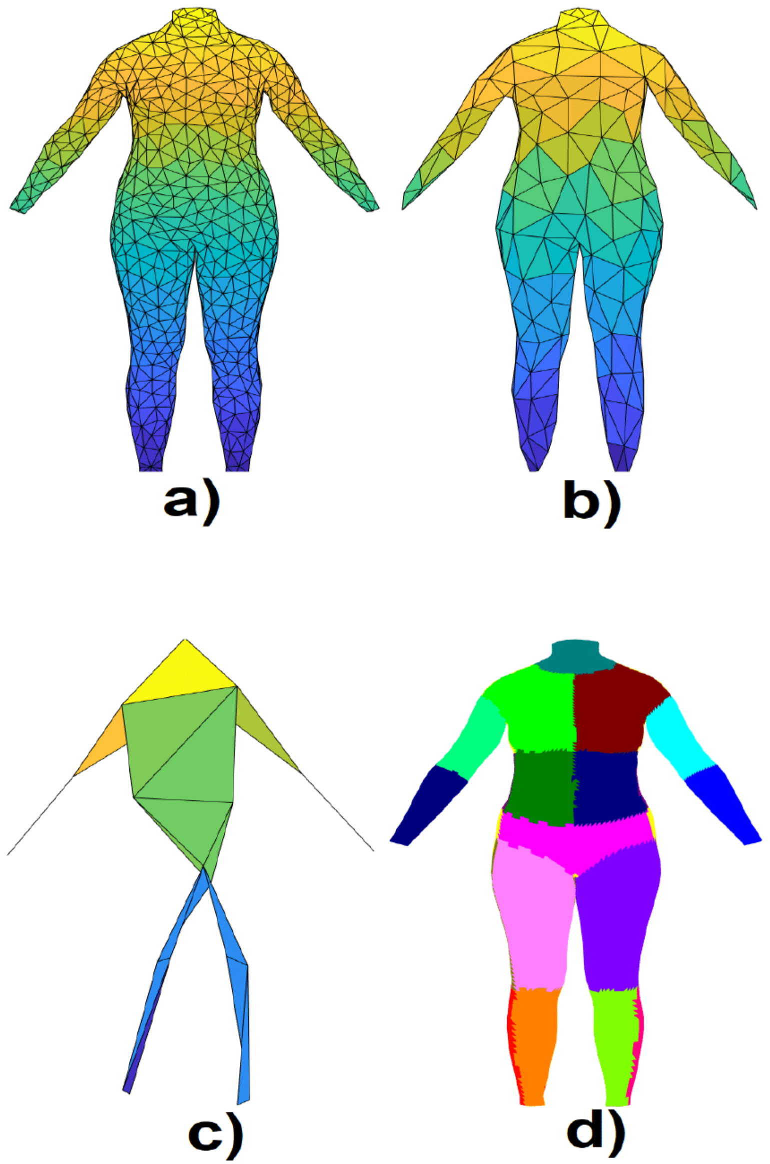
$$y = \sum_{i=0}^{K} w_i T_i(\hat{L}) x$$

- w_i: learnable weights
- $T_i(\hat{L})$: i^{th} Chebysev polynomial order
- K: node receptive field size
- $K = 1$ is used here

Overview of CLOTH3D Architecture

Pooling

- Reduce number of vertices in the graph
- Body segmented into 21 regions
- Mesh Hierarchy
  - 14475 → 3618 → 904 → 226 → 56 → 21



Network Changes

- Goal:
  - Study different architectures for 3D garment reconstruction
- Modifications to CLOTH3D architecture:
  - Pooling
  - Central difference convolution
  - Octave convolution

Network Changes

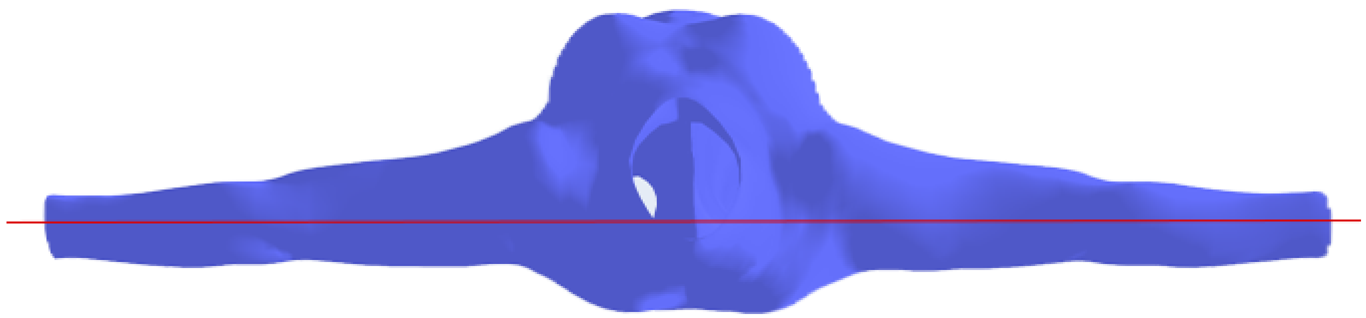
Pooling

- Objectives:
  - 6 vertices in final layer
    - 2 arms, 2 legs, front and back torso
  - Increase vertex pop. in lower arms and legs
    - These regions have high variability in cloth dynamics
  - Observe effects of alternate pooling
- Four steps:
  1. Split the vertices into front and back
  2. Recursive Spectral clustering
  3. Manually merge final layer
  4. Create the edge connections

Network Changes

Pooling

1) Split the vertices into front and back



Network Changes

Pooling

2) Spectral Clustering (SC)

- Cluster both groups (front and back) into 5 (total of 10)
  - 5 clusters: arms, legs, torso
- Recursively apply SC to each group
  - Base case: single vertex remaining
- Hierarchy:
  - $14475 \rightarrow 8648 \rightarrow 2558 \rightarrow 640 \rightarrow 160 \rightarrow 40 \rightarrow 10$

3) Manually merge final layer

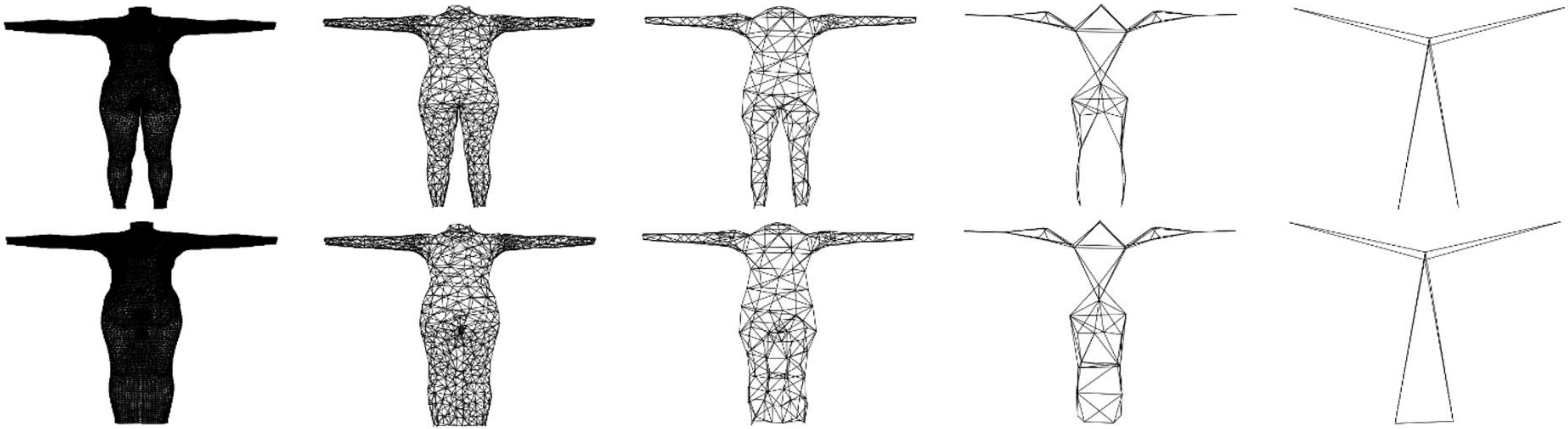
- Combine arm and leg vertices
- Final Hierarchy:
  - $14475 \rightarrow 8648 \rightarrow 2558 \rightarrow 640 \rightarrow 160 \rightarrow 40 \rightarrow 10 \rightarrow 6$

Network Changes

Pooling

4) Create edge connections

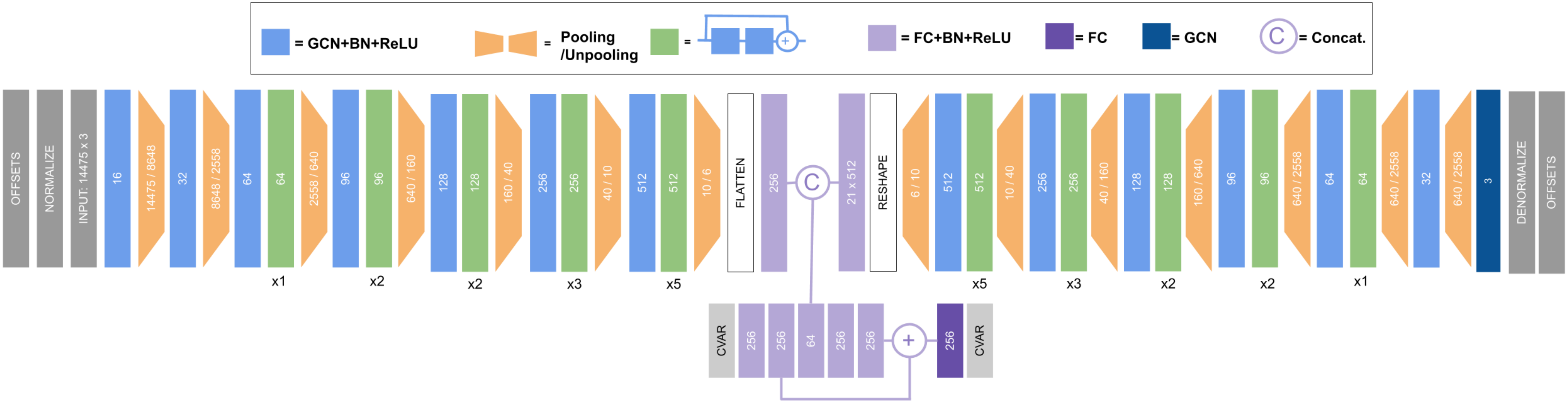
- Start with initial edges (from 14475 vertex graph)
- Consider each vertex as a group of vertices pooled together
- If an edge exists between vertices from separate groups, then an edge exists between groups



Network Changes

Pooling

- Modify architecture layers
  - CLOTH3D: 5 pooling layers
  - New Pooling: 7 pooling layers



Network Changes

Central Difference Convolution (CDC)

- Goal:
  - Enhance vanilla convolution using center oriented gradient of samples
- Vanilla convolution:
  - Weighted sum of sampled region
- Central Difference:
  - Weighted sum of difference between sampled region and center
- Vanilla + central difference = intensity + gradient learning

- CNNs

$$y(p_0) = \underbrace{\sum_{p_n \in R} w(p_n) \cdot (x(p_0 + p_n))}_{\text{Vanilla}} \underbrace{- \theta \cdot x(p_0) \cdot \sum_{p_n \in R} w(p_n)}_{\text{Central Difference}}$$

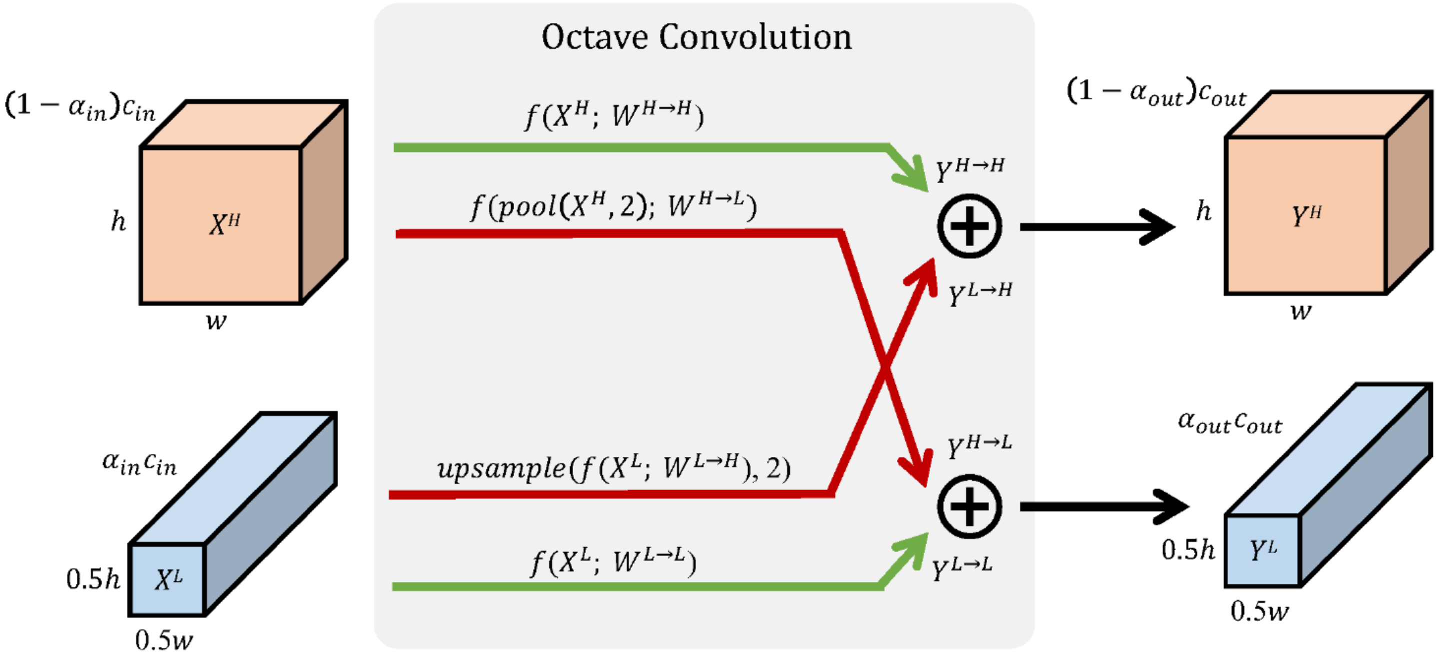
- GNNs
  - $\theta = 0.7$
  - $K = 1$

$$y = \underbrace{\sum_{i=0}^{K} w_i T_i(\hat{L}) x}_{\text{GCN}} \underbrace{- \theta \cdot x \cdot w_1}_{\text{Central Difference}}$$

Network Changes

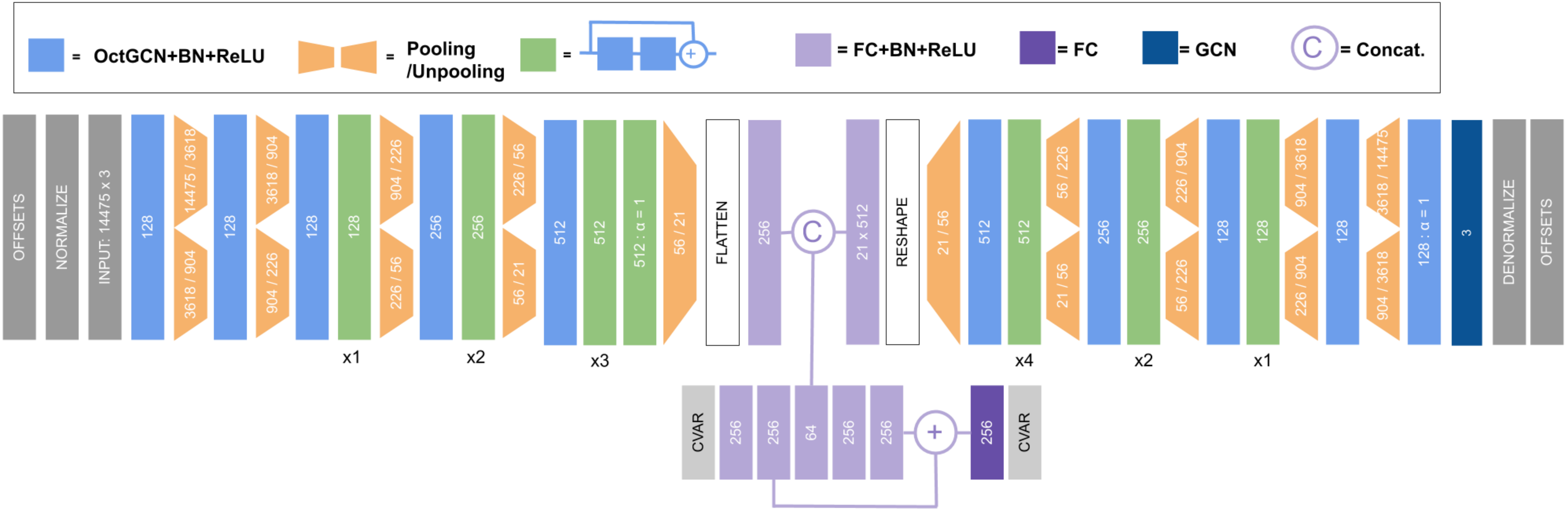
Octave Convolution

- Goal:
  - Increase efficiency by separating features into high and low components
- $\alpha = 0.875$



Network Changes

Octave Convolution



Experiments

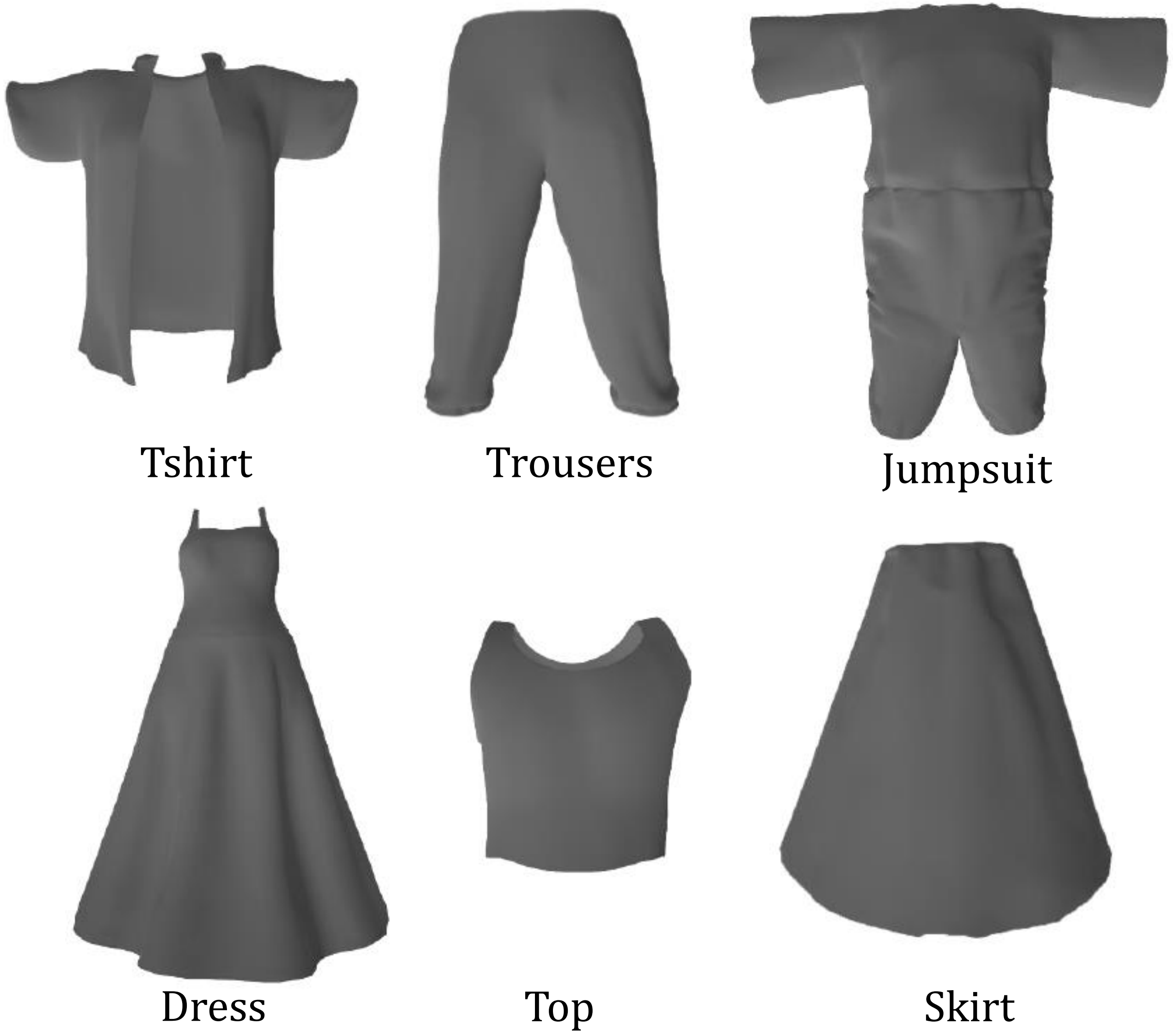
Networks

- Five networks tested:
  - Default
  - CDC
  - Pooling
  - Octave
  - CDC-Pooling
- Quantitative and Qualitative Results

Experiments

Data

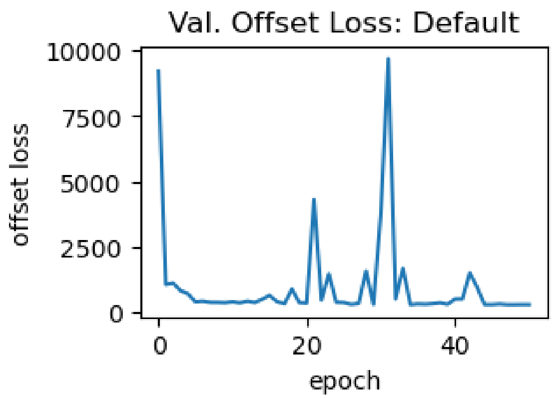
- Subset of CLOTH3D Dataset
- Training: 647 outfits
- Validation: 60 outfits
- 300 frames per outfit
- 6 Garment types
- Fabrics:
  - Silk, denim, cotton, leather

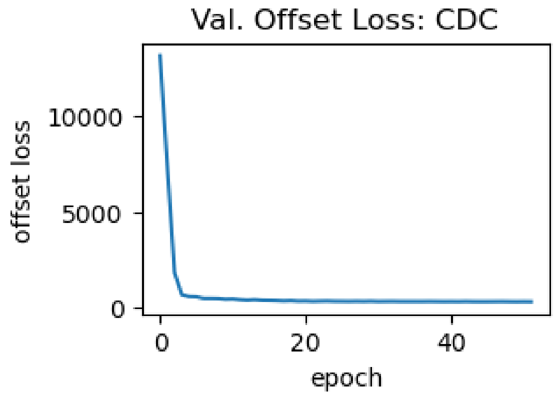


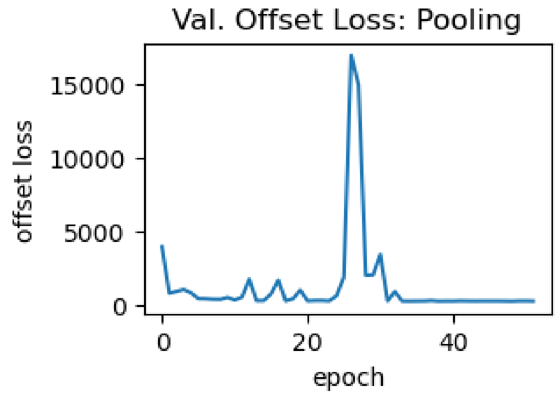
Experiments

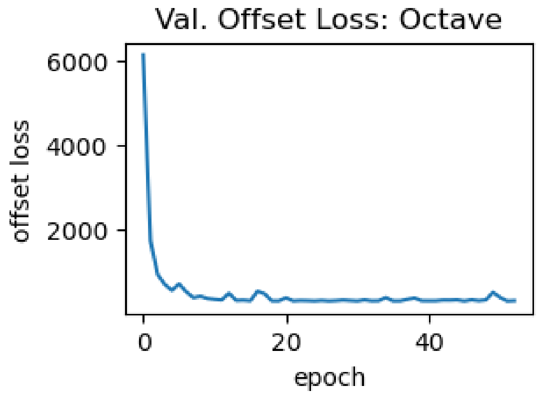
Training

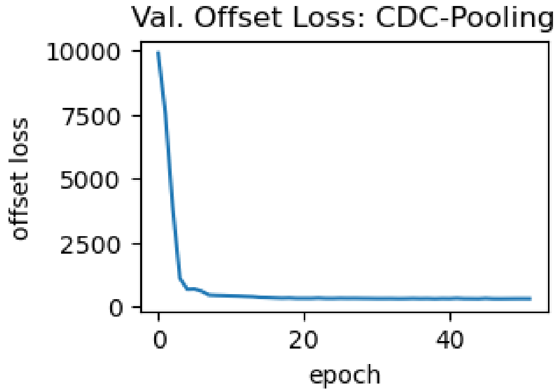
- NVIDIA GeForce GTX 1080Ti
- Epochs: 52
- LR: 0.0001
- Optimizer: Adam
- Batch Size: 16
- Incremental Training by 200 steps
- Offset Loss







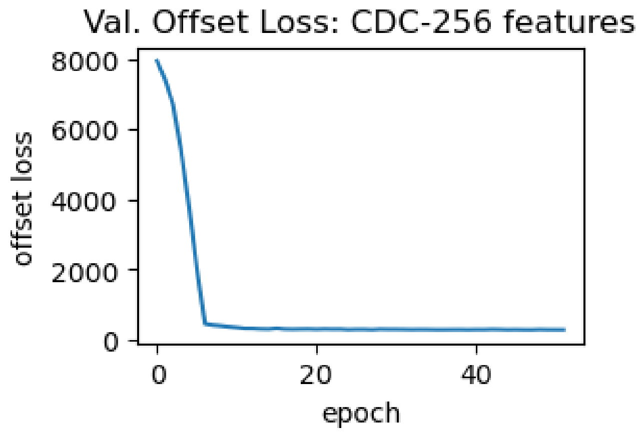


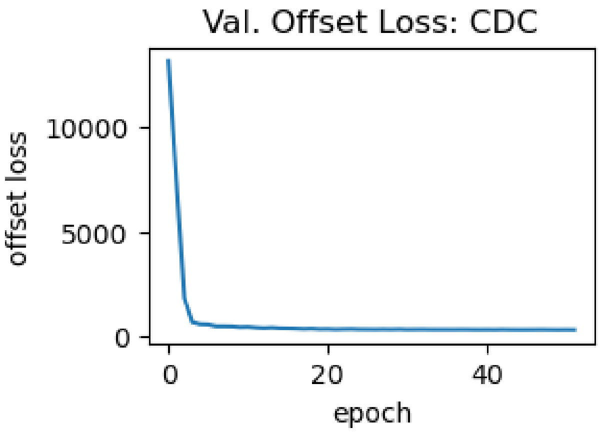


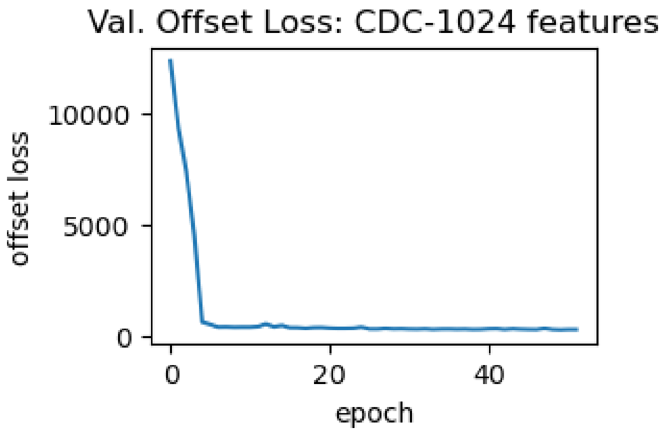
Experiments

Ablation Study

- Capacity of the network
  - CDC Network
  - 256, 512, 1024 features
- 512 has fastest convergence
- 512 lowest overall loss







Experiments

Quantitative Results

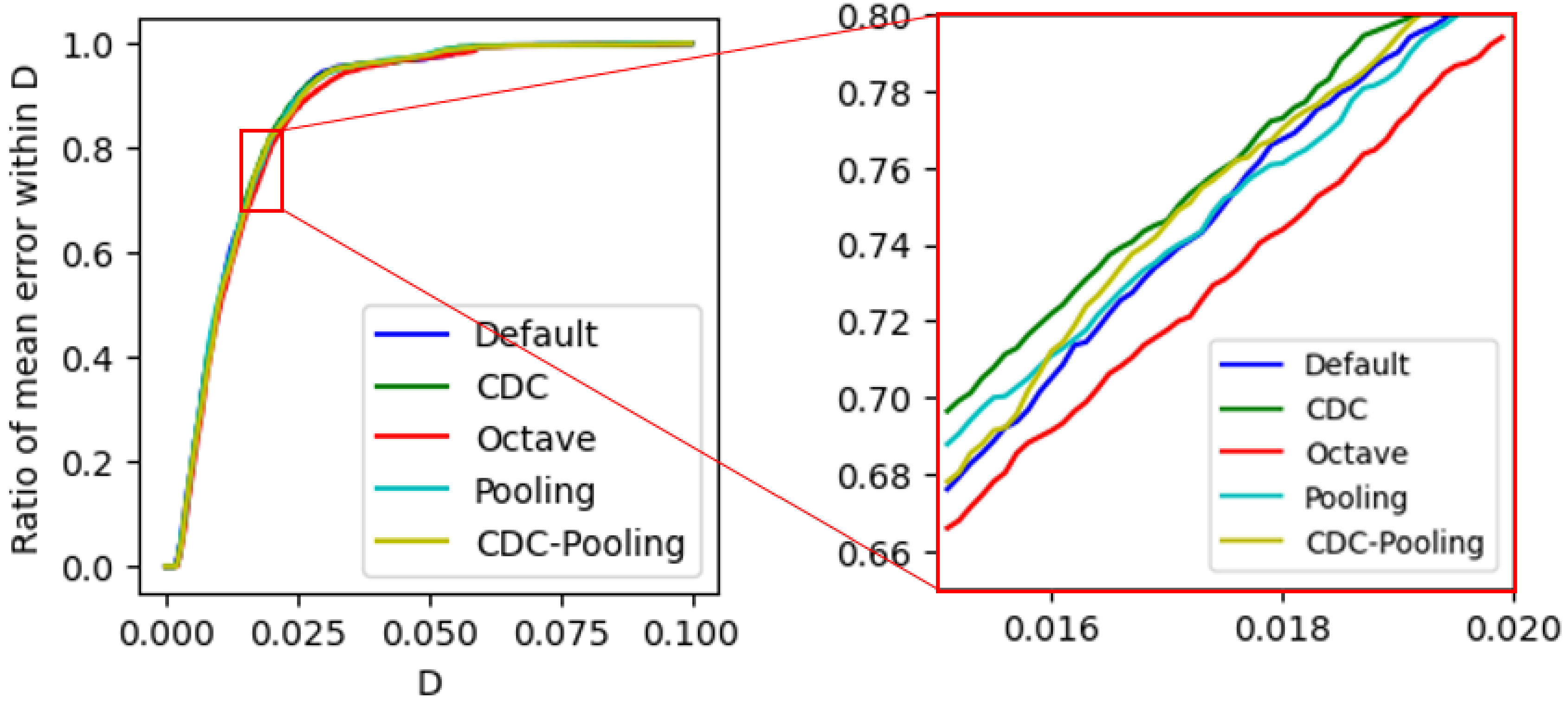
	Default	CDC	Pooling	Octave	CDC-Pool
Tshirt	0.0282	0.0281	**0.0271**	0.0293	0.0282
Trousers	0.0245	0.0254	0.0254	0.0266	**0.0241**
Jumpsuit	0.0202	0.0204	**0.0201**	0.0210	0.0202
Dress	0.0375	**0.0371**	0.0374	0.0404	0.0381
Top	0.0212	0.0211	**0.0209**	0.0219	0.0230
Skirt	0.0639	**0.0610**	0.0624	0.0650	0.0641
silk	0.0397	0.0395	**0.0394**	0.0415	0.0408
denim	0.0306	**0.0303**	0.0304	0.0316	**0.0303**
cotton	0.0246	0.0248	**0.0245**	0.0262	0.0256
leather	0.0280	0.0270	0.0277	0.0297	**0.0266**
TOTAL	0.0304	**0.0302**	**0.0302**	0.0320	0.0308

Average per vertex Euclidean error of garment vertices, for each architecture.

Experiments

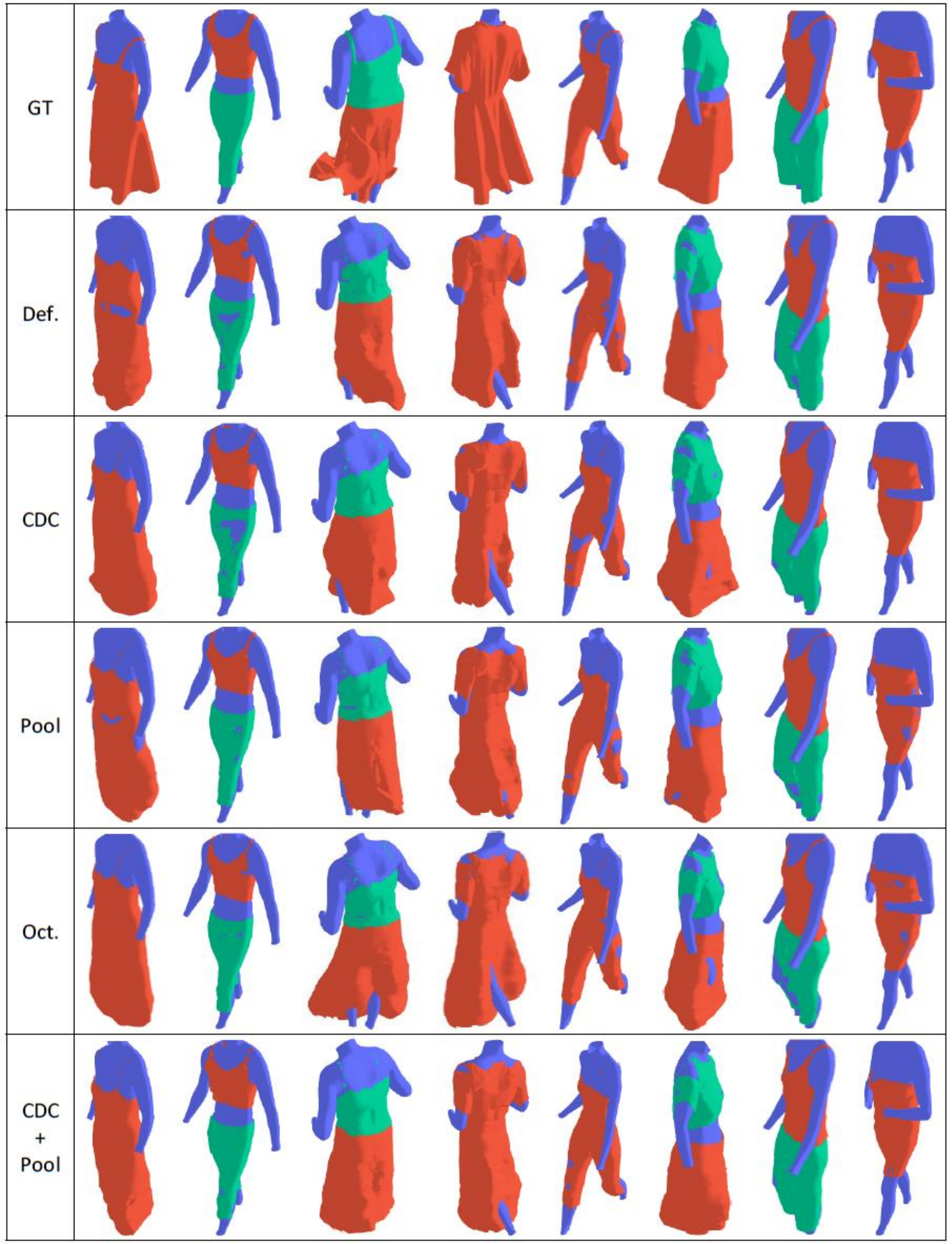
Quantitative Results

- Frame Ratio above Distance Threshold D



Experiments

Qualitative Results



Conclusions

- Study of architectures for 3D garment reconstruction
- Five networks compared
- Quantitative results:
  - CDC and Pooling modifications are positive
  - Octave modification negative
- Qualitative Results:
  - Similar results for: Tshirt, Trousers, Top, Jumpsuit
  - Pooling captured the cloth dynamics of the skirt and dress garments best
  - CDC-Pooling had best overall performance
- Future Considerations:
  - Explore and improve the pooling
  - Account for penetrations

Questions?