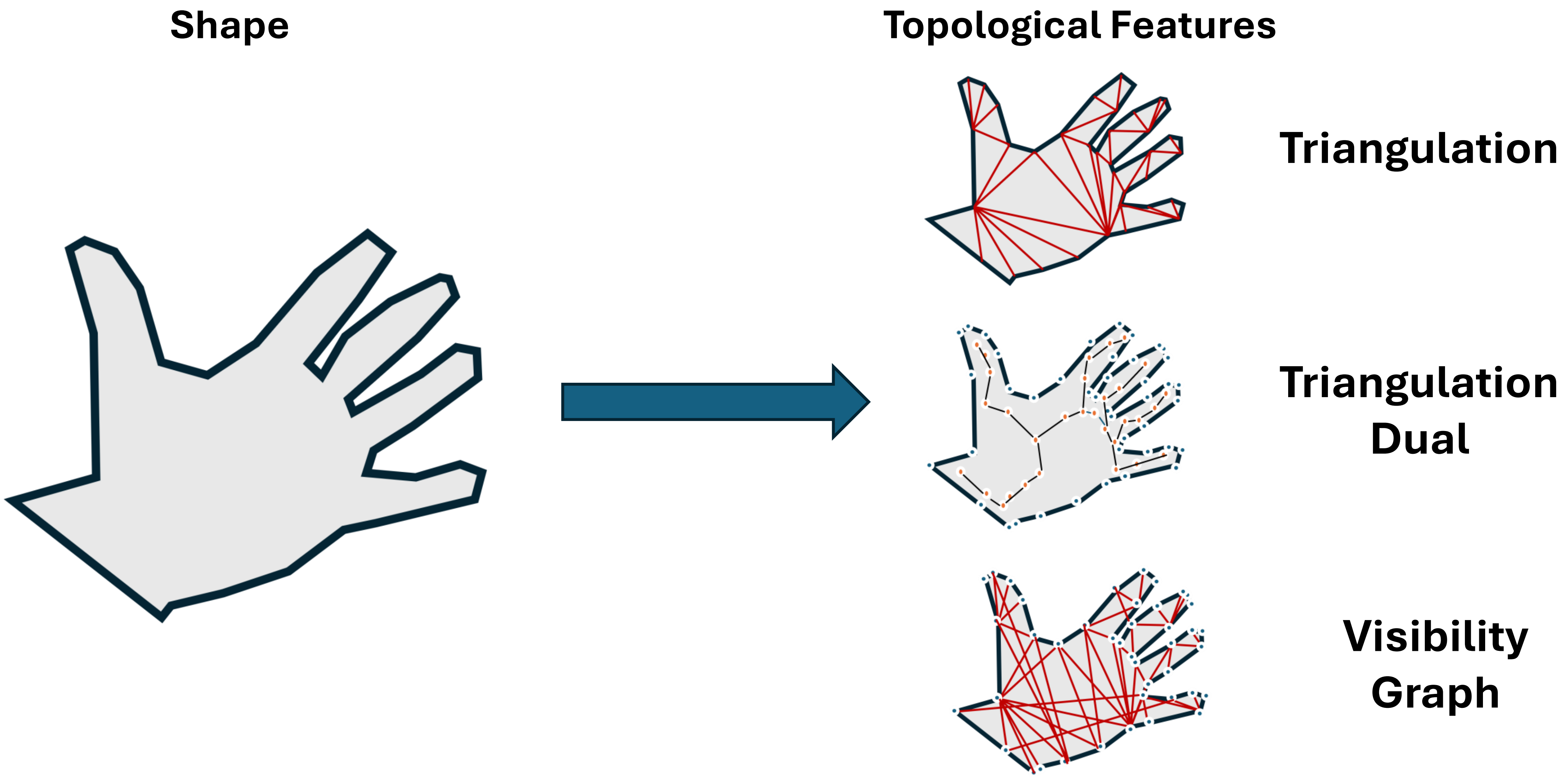
VisDiff: SDF-Guided Polygon Generation for Visibility Reconstruction, Characterization and Recognition

Rahul Mahesh[1], Jun-Jee Chao[1] and Volkan Isler[2]

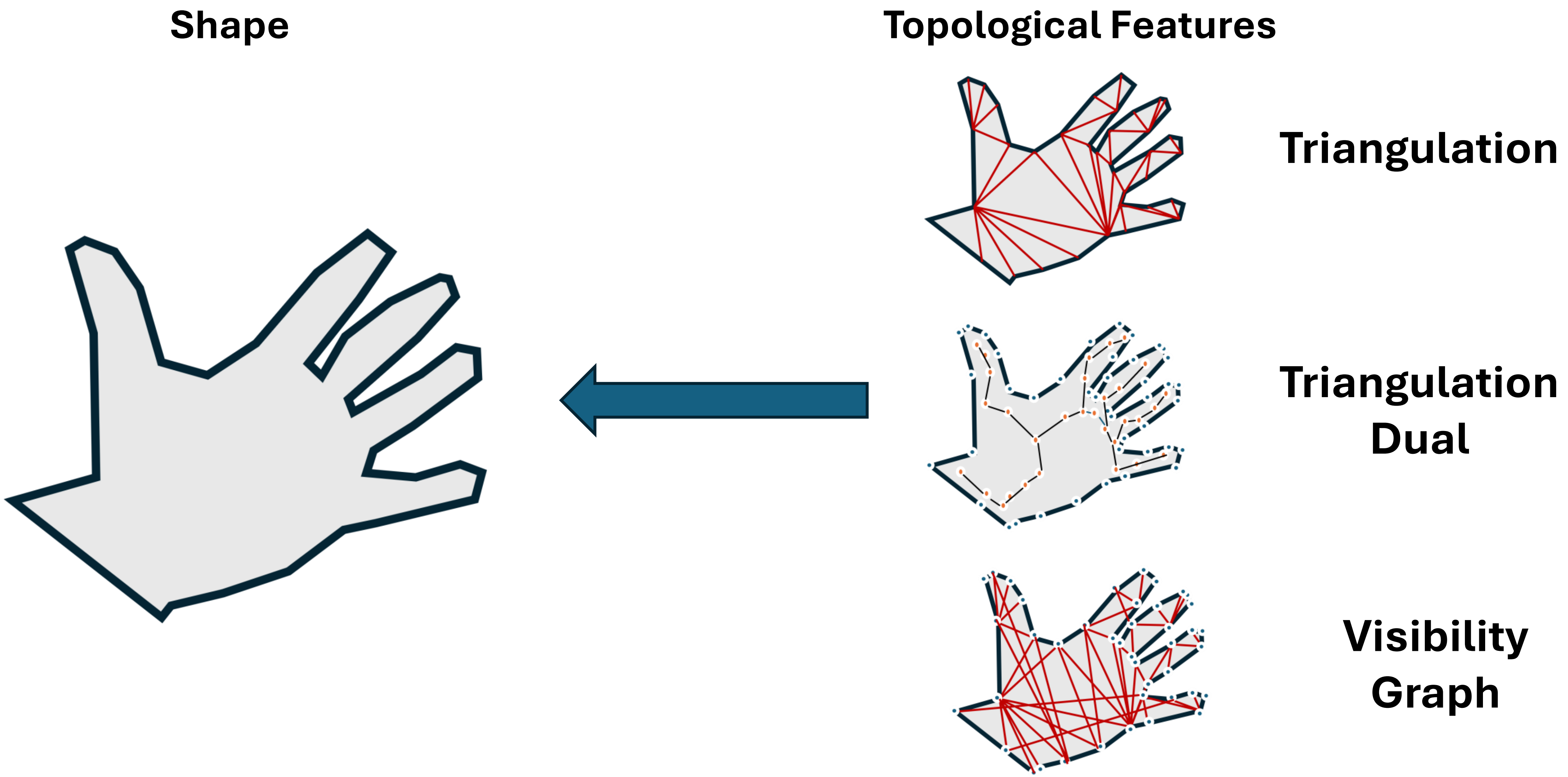
Robotics, Sensing and Networks Lab

[1]University of Minnesota and [2]The University of Texas at Austin

Motivation

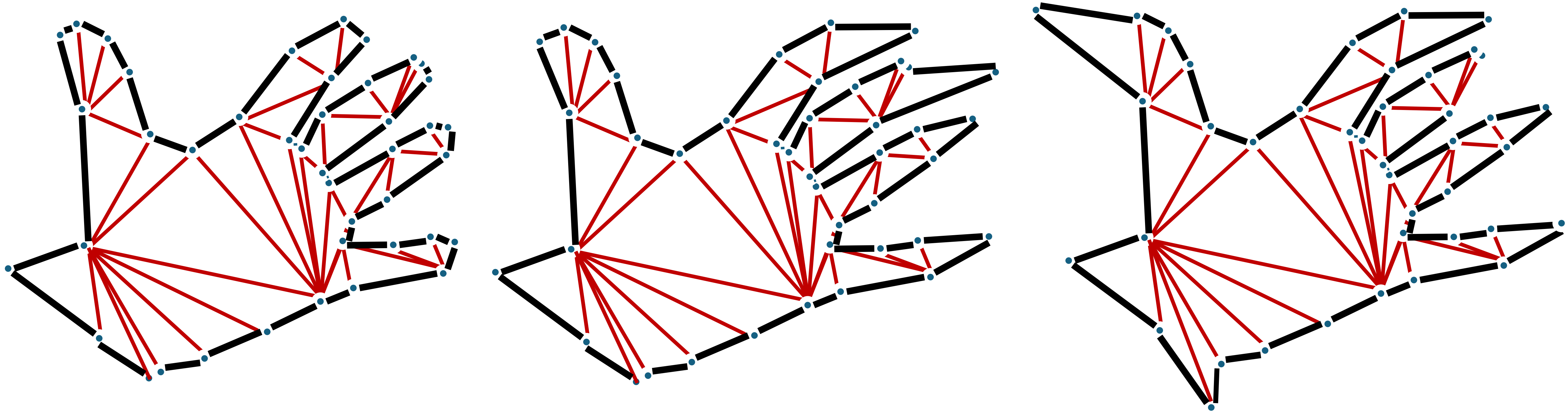

Shape
Topological Features
Triangulation
Triangulation Dual
Visibility Graph

Motivation


Shape
Topological Features
Triangulation
Triangulation Dual
Visibility Graph

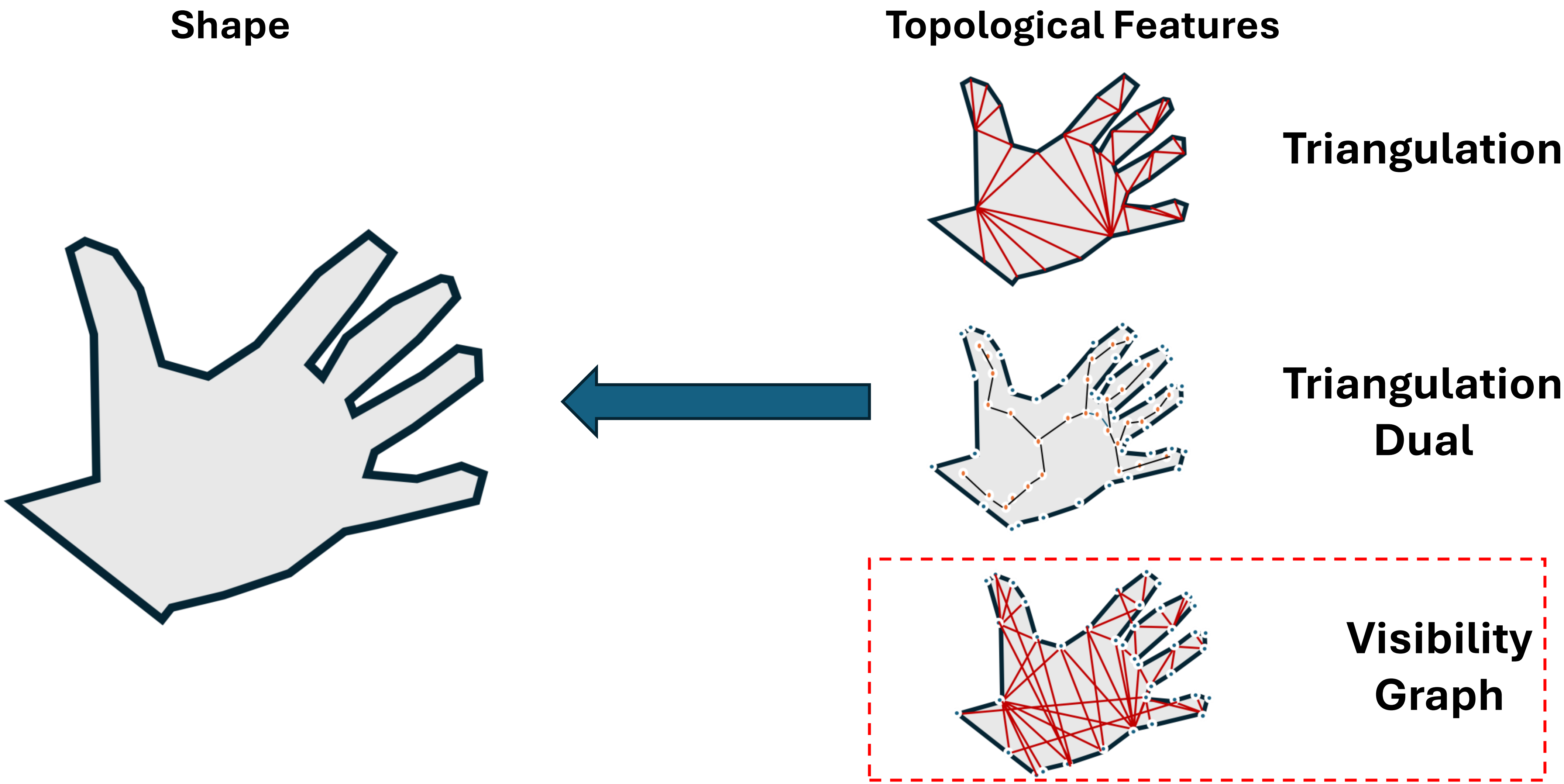
Motivation

- Multiple Polygons can exist with the same visibility graph / topological feature



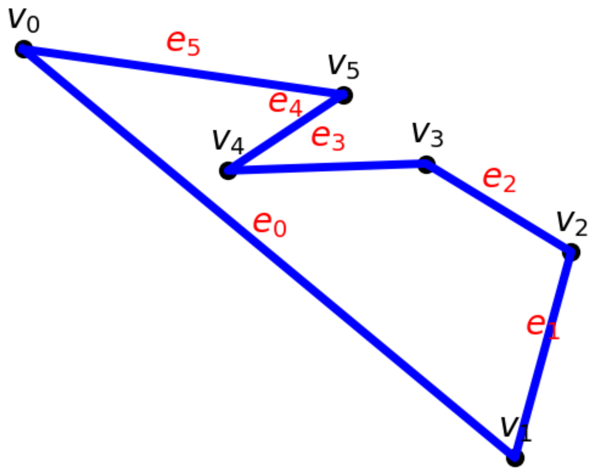
Note: Only few visibility edges are visualized

Motivation

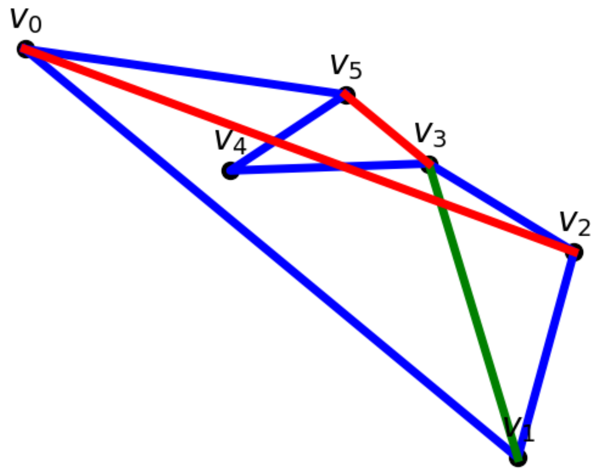

Shape
Topological Features
Triangulation
Triangulation Dual
Visibility Graph

What is a Visibility Graph?

- Two vertex locations are visible to each other if line segment joining them is contained inside the polygon

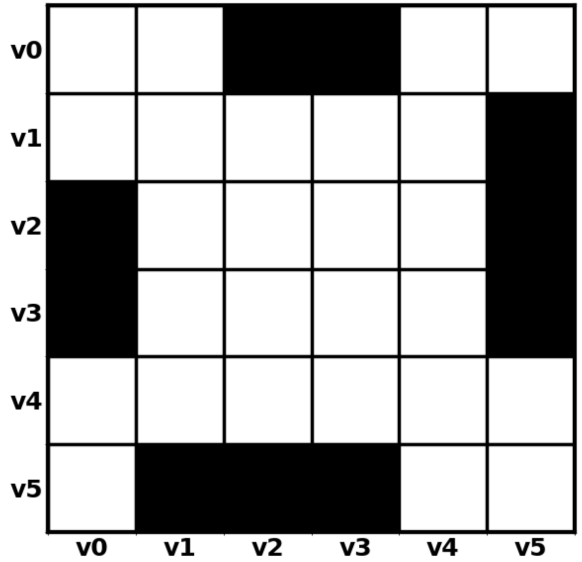


Simple Polygon



Example Visibility Edges

Green: Visible Edges
Red: Non-Visible Edges



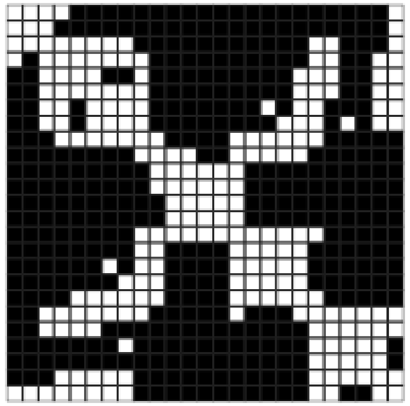
Visibility Graph

White (1): Visible Edges
Black (0): Non-Visible Edges

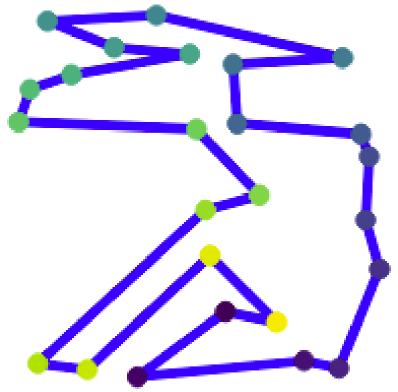
Problem Formulation:

- **Reconstruction**
  - Given a **valid visibility graph** G, generate **a** polygon P such that Vis(P) = G
- **Characterization**
  - Given a **valid visibility graph** G, generate **all** polygons P such that Vis(P) = G
- **Recognition**
  - Given an **arbitrary graph** G, determine whether there **exists** a polygon P such that Vis(P) = G
- Solved only for specific polygon classes, general cases remain open.

A graph G is ***valid*** if at least one polygon exists for it.



G Only



P

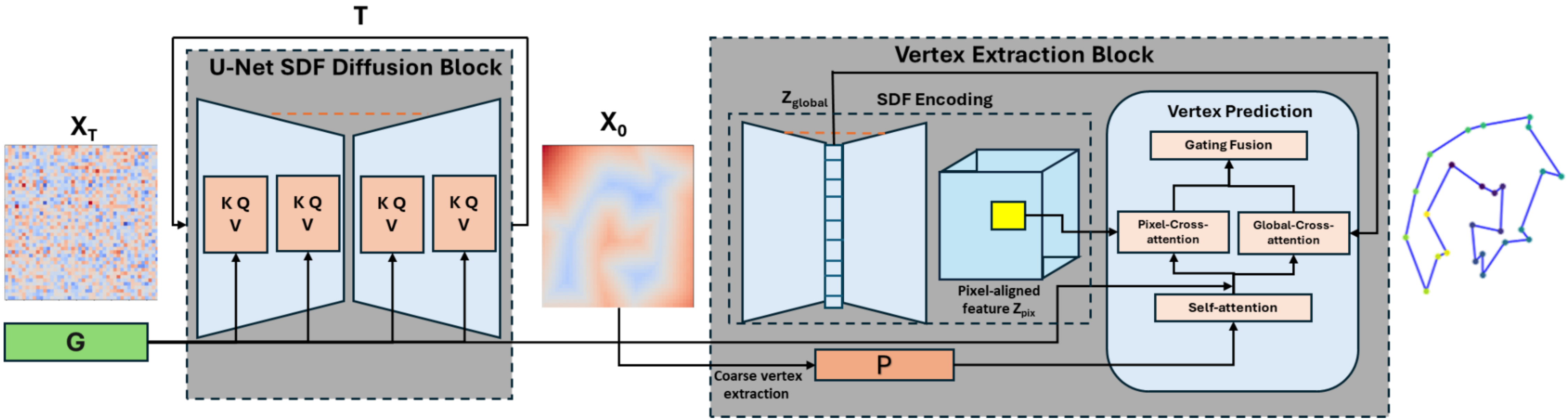
Core Contribution

- We propose **VisDiff** to solve the following–
  - **Reconstruction** - Given a **valid visibility graph** G, generate **a** polygon P such that Vis(P) = G
  - **Characterization** - Given a **valid visibility graph** G, generate **all** polygons P such that Vis(P) = G
  - **Recognition** - Given an **arbitrary graph** G, determine whether there **exists** a polygon P such that Vis(P) = G

Core Contribution

- We propose **VisDiff** to solve the following–
  - **Reconstruction** - Given a **valid visibility graph** G, generate **a** polygon P such that Vis(P) = G
  - **Characterization** - Given a **valid visibility graph** G, generate **all** polygons P such that Vis(P) = G
  - **Recognition** - Given an **arbitrary graph** G, determine whether there **exists** a polygon P such that Vis(P) = G
- Demonstrate that utilizing the SDF **enhances** the efficiency of the current learning approach in understanding visibility relationships.

VisDiff: Architecture



VisDiff: Approach

- **Reconstruction**
  - **Input:** Valid visibility graph **G**
  - **Initialize:** Single seed sampled from a Gaussian
  - **Output:** Polygon **P** with visibility graph **G'**
- **Characterization**
  - **Input:** Valid visibility graph **G**
  - **Initialize:** Multiple seeds sampled from a Gaussian
  - **Output:** Multiple Polygons **P** with visibility graphs **G'**

VisDiff: Approach

- **Reconstruction**
  - **Input:** Valid visibility graph **G**
  - **Initialize:** Single seed sampled from a Gaussian
  - **Output:** Polygon **P** with visibility graph **G'**
- **Characterization**
  - **Input:** Valid visibility graph **G**
  - **Initialize:** Multiple seeds sampled from a Gaussian
  - **Output:** Multiple Polygons **P** with visibility graphs **G'**
- **Recognition**
  - **Input: G**
  - **Initialize:** Multiple seeds sampled from a Gaussian
  - **Output:** Multiple Polygons **P** with visibility graphs **G'**
  - **Classify** valid if G – G' > T (Threshold)

Results - Quantitative Metrics

- Reconstruction and Recognition
    - Given ground truth G and G' of the predicted polygon-
        - Classification metrics
            - Accuracy
            - Precision
            - Recall
            - F1-Score
        - We primarily use F1-Score as visibility graphs can contain imbalance between visible and non-visible edges

Results - Quantitative Metrics

- Reconstruction and Recognition
  - Given ground truth G and G' of the predicted polygon-
    - Classification metrics
      - Accuracy
      - Precision
      - Recall
      - F1-Score
    - We primarily use F1-Score as visibility graphs can contain imbalance between visible and non-visible edges
- Characterization
  - Given the set of polygons **P** with with visibility graphs **G'**
    - Diversity: Average Chamfer distance between point sets of **P**
    - Coverage: Breadth-First exploration over the latent space
      - Initialize root polygon **P** using VisDiff.
      - Breadth-first exploration is then performed up to a fixed depth **d** and branching factor **b**
        - Children generated by adding scheduled noise to its parent in latent space.
      - Node expanded: **(1)** F1 greater than **T** (Same as Recognition), **(2)** Distance from previous nodes greater than $\mathbf{T_d}$
      - **Coverage:** expanded nodes / maximum possible node

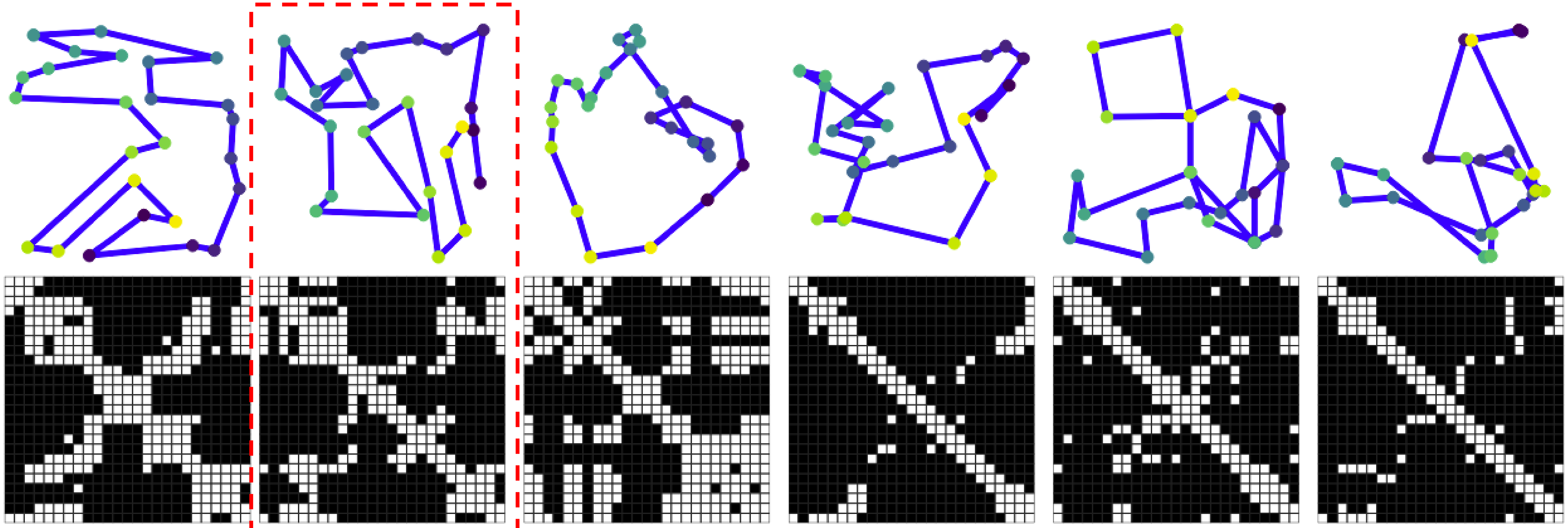
Baselines

- No existing learning method for visibility → polygon mapping
- We use generative models learning similar topological mappings as baselines-
  - **Triangulation-conditioned Mesh Generation**: MeshAnything [1], Vertex-Diffusion [2]
  - **Graph Embedding**: GNN [3]
  - **Conditional-Generation:** Conditional-VAE [4]
  - **Note:** We train the following models for visibility-polygon mapping

Quantitative Metrics - Reconstruction

	Accuracy	Precision	Recall	F1
Vertex-Diffusion [2]	0.777	0.7773	0.716	0.724
Conditional-VAE [4]	0.74	0.718	0.704	0.702
GNN [3]	0.73	0.786	0.686	0.674
MeshAnything [1]	0.7747	0.739	0.723	0.712
VisDiff	**0.924**	**0.914**	**0.911**	**0.912**

Qualitative Results- Reconstruction



GT | VisDiff F1: 0.855 | Conditional-VAE F1: 0.653 | Vertex-Diffusion F1: 0.736 | MeshAnything F1: 0.745 | GNN F1: 0.76

White: Visible Edges
Black: Non-Visible Edges

Quantitative Metrics - Characterization

- Diversity
  - Mean Chamfer Distance: **0.56**
    - N = 50
    - **High** Diversity: **20%** of 2x2 domain

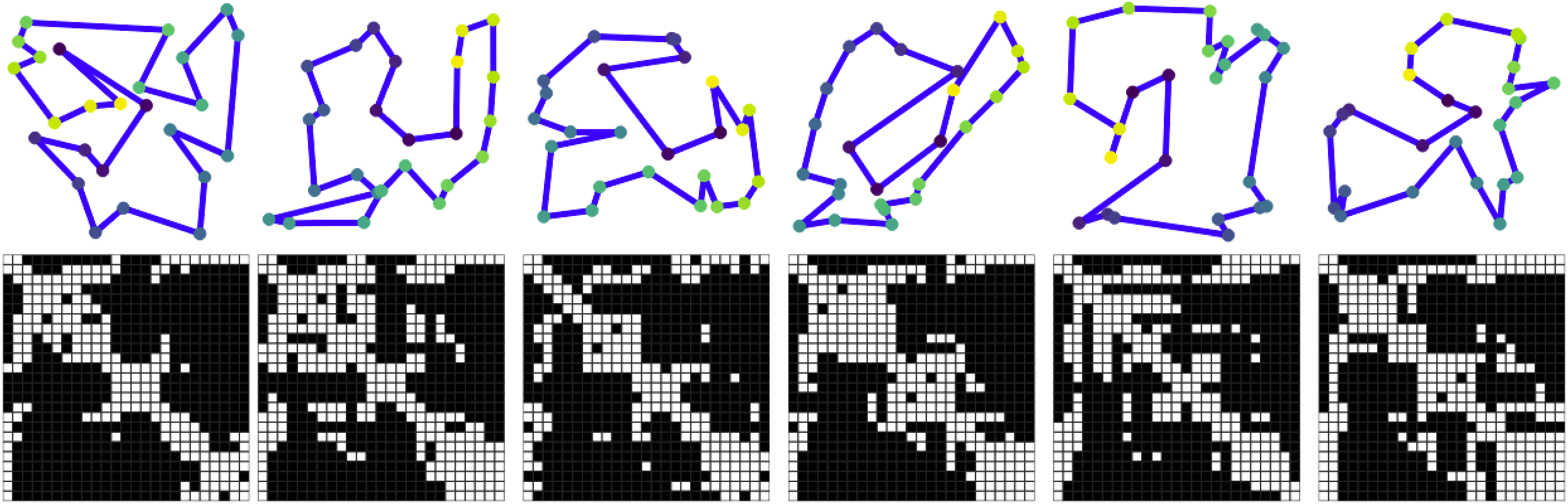
Quantitative Metrics - Characterization

- Diversity
  - Mean Chamfer Distance: **0.56**
    - N = 50
    - **High** Diversity: **20%** of 2x2 domain
- Coverage

F1 Threshold T	Depth d	Branching Factor b	Distance Threshold T_d	Coverage Metric
0.85	5	2	0.1	0.475
0.80	5	2	0.1	0.488
0.75	5	2	0.1	0.495
0.70	5	2	0.1	**0.515**

- Average: **50%** coverage metrics **~32** nodes indicating **high coverage**
  - Training: **20** Augmentations

Qualitative Results- Characterization



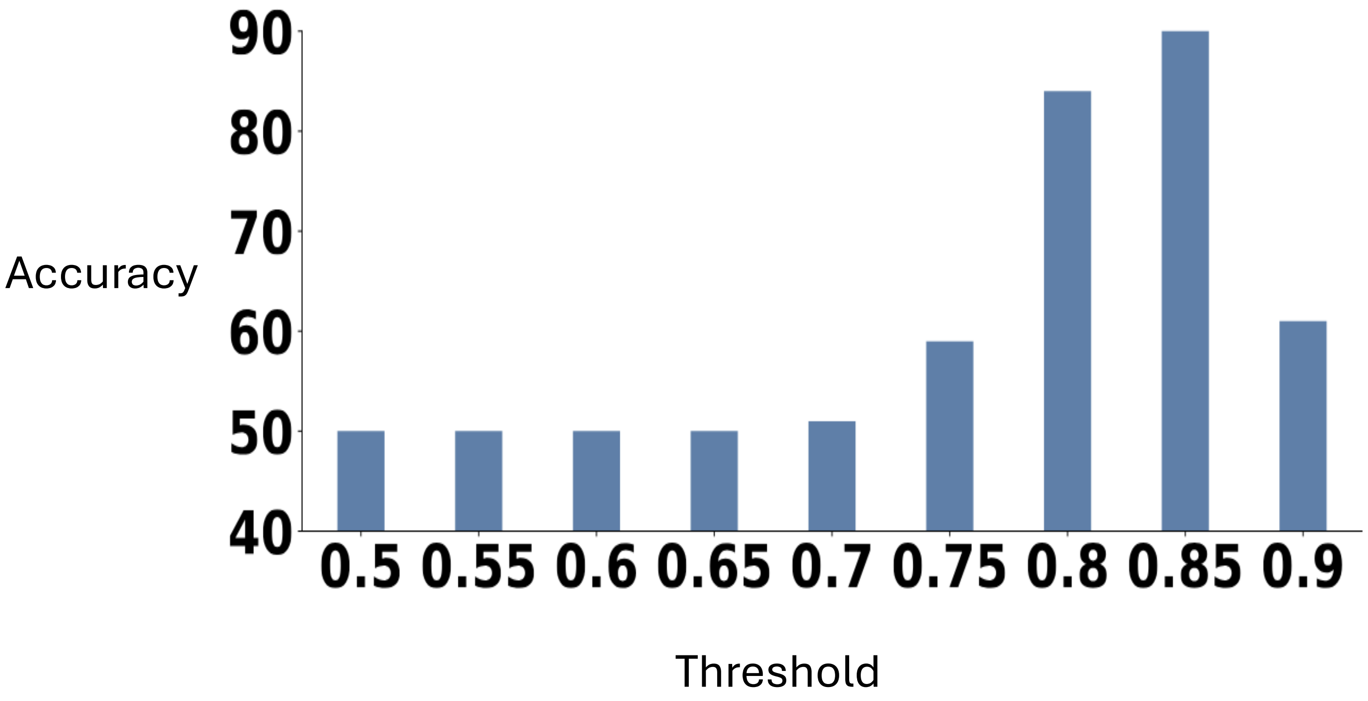
GT | F1: 0.827 | F1: 0.811 | F1: 0.827 | F1: 0.8 | F1: 0.811

White: Visible Edges
Black: Non-Visible Edges

rsn.cs.utexas.edu

Quantitative Metrics - Recognition

Polygons with **holes** are used as proxy for **non-valid visibility graphs**

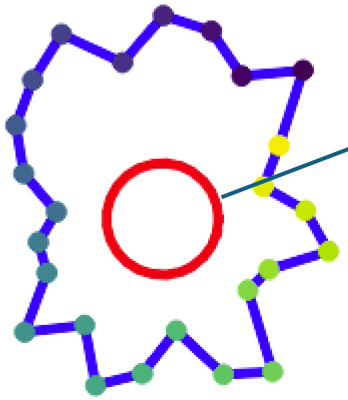


Threshold - F1 Score
Accuracy - Classification Accuracy

- **90%** Accuracy achieved with 0.85 threshold F-1 Score

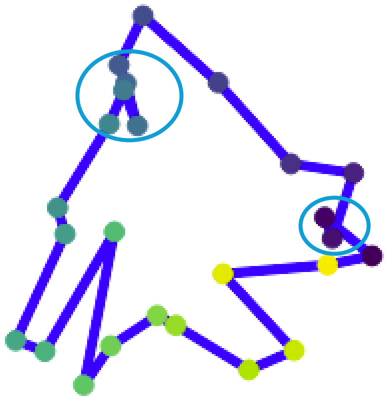
Qualitative Results - Recognition

shows crossings



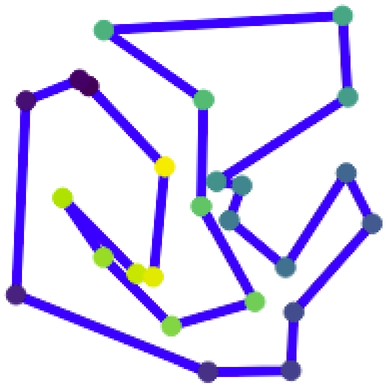
Hole

No valid polygon generated by VisDiff

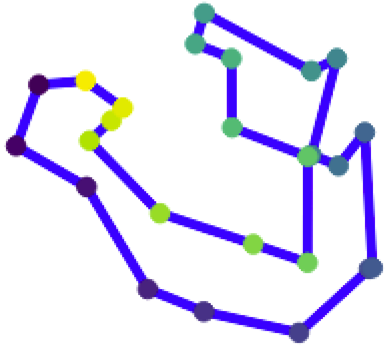


Prediction:
Non-Valid
Visibility graph

Non-Valid Visibility Graph



Valid polygon with F1 0.87 generated by VisDiff



Prediction:
Valid Visibility graph

Valid Visibility Graph

Conclusion

- Demonstrated effectiveness of having an intermediate **SDF** representation which enhances the understanding of visibility relationships
  - **VisDiff** outperforms to the state of the art approaches on the visibility reconstruction task by **26%** in terms of F-1 Score
- Showcased the capability of **VisDiff** to provide evidence for the visibility characterization and Recognition problem

Thank You!

References

1. Chen, Yiwen, Tong He, Di Huang, Weicai Ye, Sijin Chen, Jiaxiang Tang, Xin Chen et al. "Meshanything: Artist-created mesh generation with autoregressive transformers." *arXiv preprint arXiv:2406.10163* (2024).
2. Alliegro, Antonio, Yawar Siddiqui, Tatiana Tommasi, and Matthias Nießner. "Polydiff: Generating 3d polygonal meshes with diffusion models." *arXiv preprint arXiv:2312.11417* (2023).
3. Gao, Jianliang, Tengfei Lyu, Fan Xiong, Jianxin Wang, Weimao Ke, and Zhao Li. "MGNN: A multimodal graph neural network for predicting the survival of cancer patients." In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 1697-1700. 2020.
4. Harvey, William, Saeid Naderiparizi, and Frank Wood. "Conditional image generation by conditioning variational auto-encoders." *arXiv preprint arXiv:2102.12037* (2021).