Toward Edge-Level Comparison of Sewing Patterns: A Construction-Aware Edge-Role Representation

Hyungjin Son[1] and Dukyun Nam[1,*]

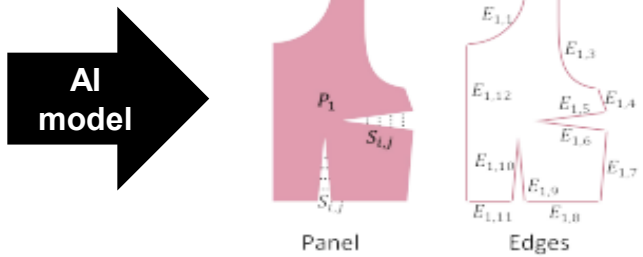
[1]High-Performance Computing Laboratory, Kyungpook National University

* Corresponding author

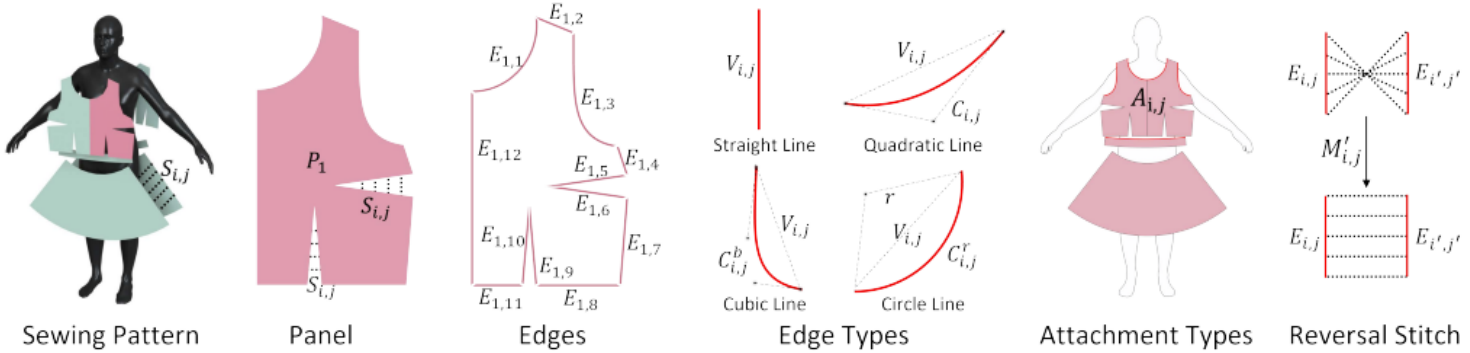
Image reference : https://www.myaiart.io/features/ai-sewing-pattern-generator/

Sewing-Pattern Generation Models

or



- A sewing-pattern generation model is an AI model that takes a sketch or text as input and generates the sewing pattern used to make the garment.



Panels, edges, edge types, attachments and reversal stitches (source: SewingLDM).

$E^f_{i,j}$ = [V, C, C^b, C^r, S, R, T, E^t, E^m, A, M′]

Symbol	Meaning
i	panel index
j	edge index within the panel
f	feature
V	vertex / edge position
C	curve parameter
C^b	Bézier curve
C^r	circular curve
S	stitch
R	panel rotation
T	panel translation
E^t	edge type
E^m	edge mask
A	attachment constraint
M′	attachment mask

Per-edge feature vector (SewingLDM).

- Generation studies encode panels, edges, stitches and per-edge attributes as structured data and predict them, much like object detection decomposes a scene into components with attributes.

The Geometry Problem

- Geometry = the edge-level numbers that decide whether a garment can be sewn: length, direction, curvature, and equal length of stitched edges.
- Current metrics (below) reward panels and edges that exist and shapes that overlap; centimetre-level edge errors stay invisible.

Metric	Definition	Source
Panel L2 (Vertex L2)	Mean vertex L2 distance between predicted and ground-truth panels (cm)	AIpparel, SewingLDM, GarmentDiffusion, SwiftTailor
Panel Accuracy	Percentage of patterns with the correct number of panels (%)	AIpparel, GarmentDiffusion, Garment Particles
Edge Accuracy	Edge-prediction accuracy within panels when the panel count is correct (%)	AIpparel, GarmentDiffusion, SwiftTailor
Panel IoU	Intersection-over-Union between generated and ground-truth panels treated as 2D polygons (panel-shape accuracy)	Garment Particles, GarmageNet

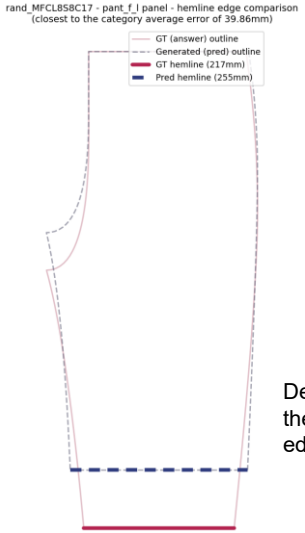
Accuracy-reporting examples, from left: SewingLDM, AIpparel, and Design2GarmentCode

Method	Accuracy (↑)	Panel L2 (↓)
SewFormer-FT†	10.3	22.4
DressCode†	0.6	31.0
AIpparel	**59.0**	**6.1**

Diversity	# Panels	$5.11_{\pm 1.66}$	$6.92_{\pm 3.10}$	$10.11_{\pm 4.42}$	$\mathbf{11.02}_{\pm 4.18}$
	# Edges	$5.48_{\pm 1.60}$	$6.84_{\pm 3.38}$	$5.79_{\pm 1.71}$	$6.24_{\pm 2.90}$
	# Stitches	$10.06_{\pm 3.24}$	$18.66_{\pm 8.64}$	$15.81_{\pm 5.91}$	$27.9_{\pm 9.83}$

Method	Panel L2(↓)	Panel Acc(↑)	Edge Acc(↑)	Rot L2(↓)	Transl L2(↓)	Stitch Acc(↑)	Failure Rate(↓)
SewFormer*	12.3	79.4	44.7	.0400	4.5	2.8	4.3%
AE (Ours)	**0.64**	**99.8**	**88.5**	**.0004**	**0.6**	**90.8**	**0**
SewingLDM	**3.13**	**97.8**	**82.7**	**.0043**	**1.2**	**84.2**	**0**

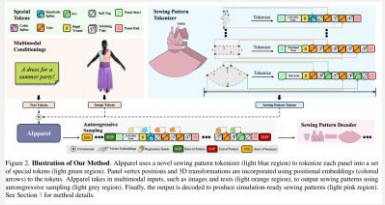
- **SewingLDM (2025)** and later work extend pattern expressiveness and condition on complex structures and multimodal inputs.
- Structural elements (panel and edge counts) are now predicted almost perfectly, but real-size geometry of each edge and vertex remains limited.
- **SwiftTailor (2026):** “Sewing patterns require centimeter-level precision … a significant challenge for neural networks that only provide statistical approximations.”
- **SewFusion (2026)** improves Vertex L2, yet edge length, direction, curvature and tangency, which decide production, are still rarely evaluated.
- → **Geometry has to be measured explicitly, edge by edge.**



Design2GarmentCode closely matches the ground-truth shape, yet visible edge-length discrepancies remain.

Prior Work: New Models, Representations and Same Geometry Gap

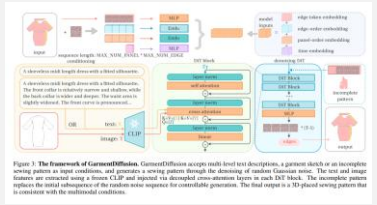
Research has mainly changed model architectures and data representations; edge-level geometry is still left behind.



LLM / VLM fine-tuning

Patterns as tokens or code for pretrained multimodal models

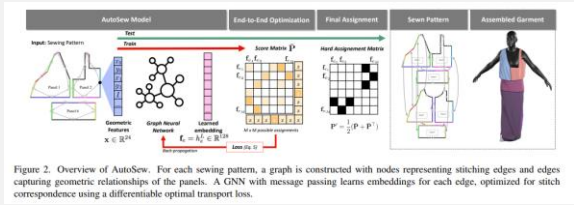
AIpparel · ChatGarment · GarmentGPT



Diffusion / flow

Generation in latent space, or edge tokens in parallel

SewingLDM · GarmentDiffusion · Cao et al.

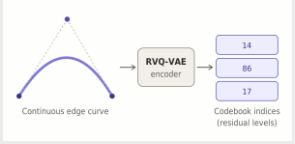


Graph / hybrid

Explicit edge relations, or discrete structure split from regression

AutoSew · SwiftTailor

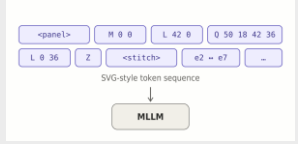
Representation changes (proposed to fix geometric errors)



GarmentGPT

Edge geometry as discrete codebook indices

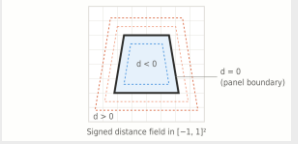
Why: VLMs are weak at regressing raw coordinates



GenPattern

SVG-style token sequence for an MLLM

Why: earlier tokenizations lose pattern structure



Cao et al.

Panels as distance fields; stitches in a separate module

Why: autoregressive generation accumulates error

95.62%

GarmentGPT panel accuracy

81.84%

GarmentGPT stitch accuracy

81.0%

Cao et al. draping success

69.0%

Cao et al. draping success

Geometry Errors Cannot Yet Be Located

Losses, metrics and data treat every edge alike, so a geometry error cannot be traced to a specific edge.

Sewformer · SIGGRAPH Asia '23 — **Role-aware: No**

$$\mathcal{L} = \lambda_1 \mathcal{L}_{\text{panel}} + \lambda_2 \mathcal{L}_{\text{stitch}} + \lambda_3 \mathcal{L}_{\text{SMPL}}$$

$$\mathcal{L}_{\text{panel}} = \mathcal{L}_{\text{shape}} + \mathcal{L}_{\text{loop}} + \mathcal{L}_{\text{RT}}$$

Panel shape, loop closure, rotation/translation, stitch and body terms.

Alpparel · CVPR '25 — **Role-aware: No**

$$\mathcal{L} = \text{CE}(\text{tokens}) + \lambda \sum_e \|\hat{v}_e - v_e\|_2 + \lambda \sum_R \|\hat{R} - R\|_2$$

$$\lambda = 0.1$$

Token cross-entropy plus one weighted L2 over all edge parameters.

Cao et al. · SIGGRAPH '26 — **Role-aware: No**

$$\mathcal{L}_{\text{panel}} = \mathcal{L}_{\text{sdf}} + \mathcal{L}_{\text{udf}} + \lambda_{\text{KL}} \mathcal{L}_{\text{KL}}$$

$$\mathcal{L}_{\text{stitch}} = \sum_i \text{BCE}(H_{\text{sew}}(f_e^i),\ l_{\text{sew}}^i) + \sum_{i,j} \text{BCE}(\sigma(f_c^{ij}),\ l_{\text{pair}}^{ij})$$

Signed / unsigned distance fields; BCE on sewing-edge and edge-pair labels.

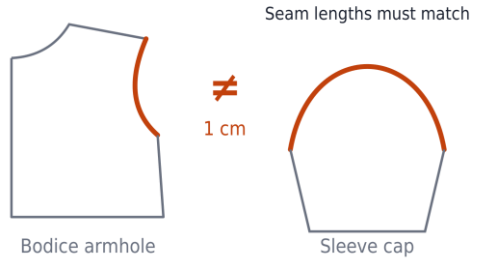
SwiftTailor · arXiv '26 — **Role-aware: No**

$$\mathcal{L} = \lambda_{\text{reg}} \mathcal{L}_{\text{reg}} + \lambda_{\text{stitch}} \text{CD}(V_A, V_B) + \lambda_{\text{norm}} \mathcal{L}_{\text{norm}}$$

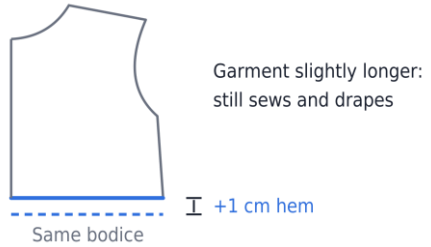
$$\mathcal{L}_{\text{reg}} = \sum_{i,j} w_{ij} |\hat{G}_{ij} - G_{ij}|$$

Weighted pixel regression, Chamfer distance between sewn edges, normals.

Today: metrics report one average (Panel L2, accuracy) and losses one sum over edges, so which edge is wrong, and by how much, cannot be read. **A role on every edge is needed first.**



Armhole: must match the sleeve cap; 1 cm off cannot be sewn.



Hemline: 1 cm longer still sews and drapes.

Same 1 cm, different consequence.

Problem Statement and Research Plan

Existing data

Edges are ordered lists without roles: a predicted edge has no defined counterpart in the ground truth.

Hidden error

Edge-level geometry errors exist but cannot be measured consistently; IoU, Panel L2 and counts average them away.

SewRole

Every edge gets a role and a fixed slot (panel, role, index), so the same edge can be compared for any model.

Goal: Show how errors that were not captured by existing evaluation metrics can be quantitatively measured using the proposed dataset.

1 Lossless conversion

Independent check against the original files: nothing lost, curve error ≤ 0.1 mm.

2 Re-score published models

Aipparel on one test set: classic metrics vs role-level metrics.

3 Known-error test

Lengthen one stitched edge of a correct pattern by a known δ: which metric reports it?

Source Data: GarmentCodeData and Zenodo 3D Garments

GarmentCodeData v2

Korosteleva et al., ECCV 2024

- Generated by GarmentCode parametric pattern programs; each sample has a pattern JSON (panels with 3D placement, vertices, curved edges, stitches), design parameters and a 3D drape.
- Used: 54,663 garments, all GCD garments merged into one set (33,678 were previously only in the training copy).
- Top+Skirt 20,218 · Top 14,125 · Skirt 12,811 · Trousers 4,156 · Top+Trousers 3,353.
- Native format of SewingLDM, AIpparel and Design2GarmentCode; rich parts (collars, cuffs, hoods, gores, circular skirts).

600,174	4.26 M	1.75 M
panels	edges	stitches

11.0 panels / garment · 7.1 edges / panel · 26% curved edges

3D Garments with Sewing Patterns

Korosteleva & Lee, NeurIPS 2021 (download from Zenodo)

- Same JSON family, generated from parametric garment templates with randomized design parameters and simulated in 3D.
- Used: 22,450 garments.
- Dress 5,150 · Jacket 4,900 · T-shirt 4,100 · Skirt 3,800 · Trousers 2,500 · Jumpsuit 2,000.
- A second design source: tests the format beyond one generator and adds dresses, jackets and jumpsuits.

124,700	716 K	252 K
panels	edges	stitches

5.6 panels / garment · 5.7 edges / panel · 28% curved edges

Total: 77,113 garments, 724,874 panels, 4.98 M edges, 2.00 M stitches; **77,051 usable** (62 excluded: 61 duplicates, 1 degenerate panel, each kept with its reason). Every garment also has front/back renders, sketches and a caption used as model input.

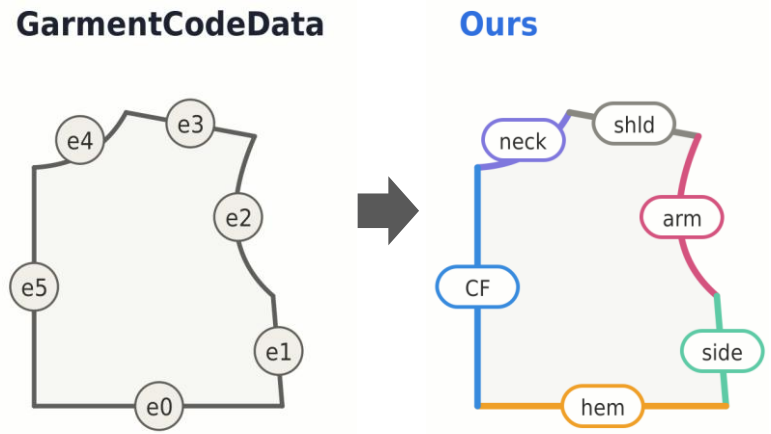
Applied Standards (1/3): Edge Roles Grounded in ISO Standards

Every edge gets one of 37 roles (25 stitched + 12 free), named after where it sits on the body, so each edge can be measured on its own.

Body region	Stitched roles (25)	Free roles (12)	ISO 8559-1 anchor
Neck, collar	neckline_seam · collar_attachment · center_seam	neckline_opening · collar_edge	neck base girth
Shoulder, armhole	shoulder_seam · armhole	armhole_edge	shoulder length, armscye girth
Sleeve, cuff	sleeve_cap · sleeve_top_seam · sleeve_underarm_seam · cuff_attachment · cuff_seam · flare_attachment	sleeve_hem · cuff_edge	arm length, wrist girth
Torso	side_seam · center_front_seam · center_back_seam · dart_leg	front_opening · back_opening	chest girth, centre front / back
Waist	waistline_seam · waistband_upper_seam · waistband_lower_seam	waistline_opening	waist girth
Skirt	panel_seam · tier_seam	hemline · slit_edge · opening	hip girth
Trousers	crotch_seam · inseam · outseam	(hemline)	crotch length, inside / outside leg length
Hood	hood_attachment · hood_center_seam	hood_edge	head girth

Columns 2–3 follow ISO 4916 (is the edge a seam or a finished edge?); column 4 follows ISO 8559-1 (where on the body is it?).

From edge numbers to edge roles



- **Before:** “e2” is only a position in a list; in another garment e2 may be a side seam.
- **After:** the same edge is “arm” in every garment, so it can be compared across models.

ISO 8559-1:2017

Size designation of clothes

Part 1: Anthropometric definitions for body measurement

Read sample

Published (Edition 1, 2017)

This publication was last reviewed and confirmed in 2026. Therefore this version remains current.

ISO 8559-1:2017 → role names

Body-measurement definitions name each role: armhole ↔ armscye girth, neckline ↔ neck base girth, waistline ↔ waist girth, inseam ↔ inside leg length.

ISO 4916:1991

Textiles — Seam types — Classification and terminology

Read sample

Published (Edition 2, 1991)

This publication was last reviewed and confirmed in 2022. Therefore this version remains current.

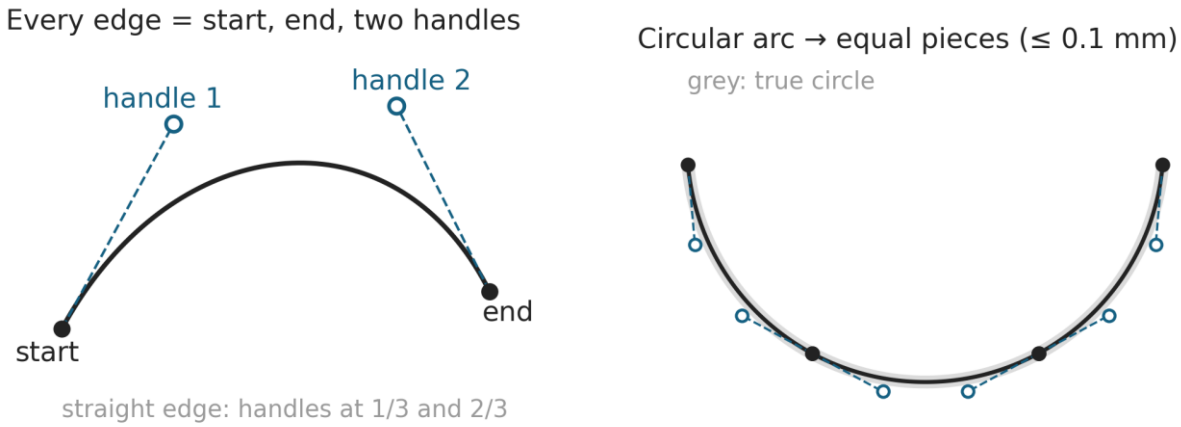
ISO 4916:1991 → stitched vs free

Seams join pieces → 25 stitched roles; edge neatening (class 6) finishes one edge → 12 free roles (hems, openings).

Applied Standards (2/3): One Format for Both Datasets

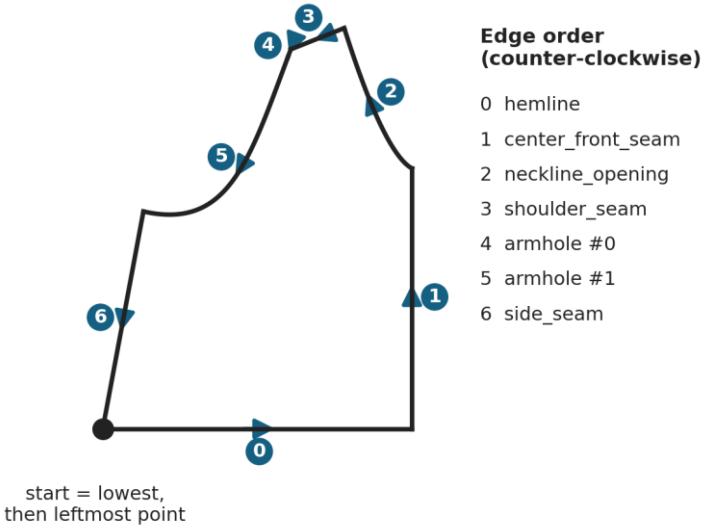
GCD and Zenodo write curves and loop direction differently; one curve form and one order rule let both share every slot.

Curves: one cubic Bézier per edge



- Every edge is stored as a start point, an end point and two handles.
- Lines and quadratic curves convert exactly; a straight edge keeps its handles at 1/3 and 2/3.
- Circular arcs are split into equal pieces, each within 0.1 mm of the circle.

Order: start point, then counter-clockwise



- Edge order keeps the inside/outside (normal) and stitch directions consistent, so the same edge always gets the same position.
- The loop starts at the lowest start point (leftmost on a tie) and runs counter-clockwise; clockwise panels are reversed.
- When an arc is split for curve accuracy, its pieces fill the original edge's place in sequence, later edges shift back, and the start point stays unchanged.

Applied Standards (3/3): Feature Dimensions

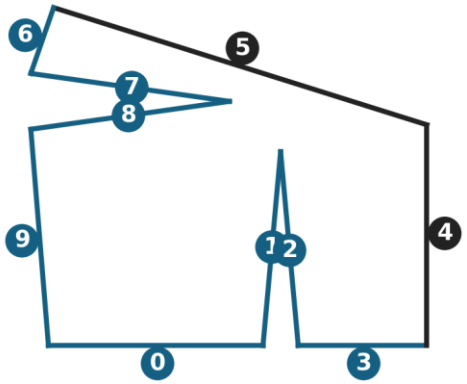
Tensor (B, 75, 63, 12)

Axis	Size	Meaning
Batch (B)	B	garments processed together
Panel slots (p)	75	13 panel kinds × left/right × front/back; repeats ordered top → bottom
Edge slots (k)	63	fixed (role, index) per panel kind
Channels	12	per-edge feature vector (right)

Why 63 edge slots for 37 roles?

A role can occur several times in one panel; each occurrence gets its own slot (role, index). The edge axis fits the largest kind: a skirt has 9 roles but needs 63 slots (panel_seam ×12, waistline_seam ×10, …).

Front bodice: 10 edges, 5 roles



edge → slot (role, index)

0 waistline_seam #0
1 dart_leg #0
2 dart_leg #1
3 waistline_seam #1
4 center_front_seam #0
5 neckline_opening #0
6 side_seam #0
7 dart_leg #2
8 dart_leg #3
9 side_seam #1

teal = role repeated in the panel

$$F_{p,k} = [x, y, c_1, c_2, s, P, K, d, o, m] \in \mathbb{R}^{12}$$

Symbol	Channels	Meaning
x, y	2	edge end point (cm, panel-local; origin = outline bbox minimum)
c_1	2	start-side Bézier handle, relative to the edge start
c_2	2	end-side Bézier handle, relative to the edge end
s	1	straight-line flag
P	1	stitch partner panel slot (−1 = free edge)
K	1	stitch partner edge slot
d	1	same stitch direction (after counter-clockwise ordering)
o	1	position in the counter-clockwise loop
m	1	existence mask (0 = empty slot, ignored in the loss)

The role is not a channel: it is the slot index k, so every model writes the same edge to the same place.

Evaluation: Standardization Introduces No Serious Error

Before showing what becomes measurable, we compare every SewRole file with its original specification to confirm that standardization itself introduces no serious error (independent checker, no code shared with the converter).

Check	GCD (54,663)	Zenodo (22,450)
Panels lost	0	0
Edges lost	0	0
Stitches lost	0	0
Vertex L2, mean	0.00001 mm	0.00002 mm
Curve error, mean	0.001 mm	0.00001 mm

- Nothing is lost: every panel, edge and stitch of the original files is kept, and vertices stay in place (Vertex L2 ≈ 0).
- The only difference is curve error. It comes from writing circular arcs as cubic Bézier pieces, and the arc splitting caps every edge at 0.1 mm.

Why a 0.1 mm curve error is acceptable

- Garment sizes are graded in steps of about 3–5 cm in chest and waist girth (KS K 0050; Kang, 2012), and the generators we evaluate still miss edge lengths by 35–85 mm on average.
- A 0.1 mm deviation is therefore hundreds of times smaller than one size step and than current model error, so it cannot change a benchmark result.
- The 0.1 mm limit is not a rough estimate. The error of approximating a circular arc with a cubic curve has an exact formula (Dokken et al., 1990), so the limit holds for every arc in the dataset.

Benchmark: Structure Is Right, Edges Are Not (Alpparel)

Alpparel (CVPR 2025), published checkpoint without retraining, scored on the GCD test set (843 garments) with SewRole slots.

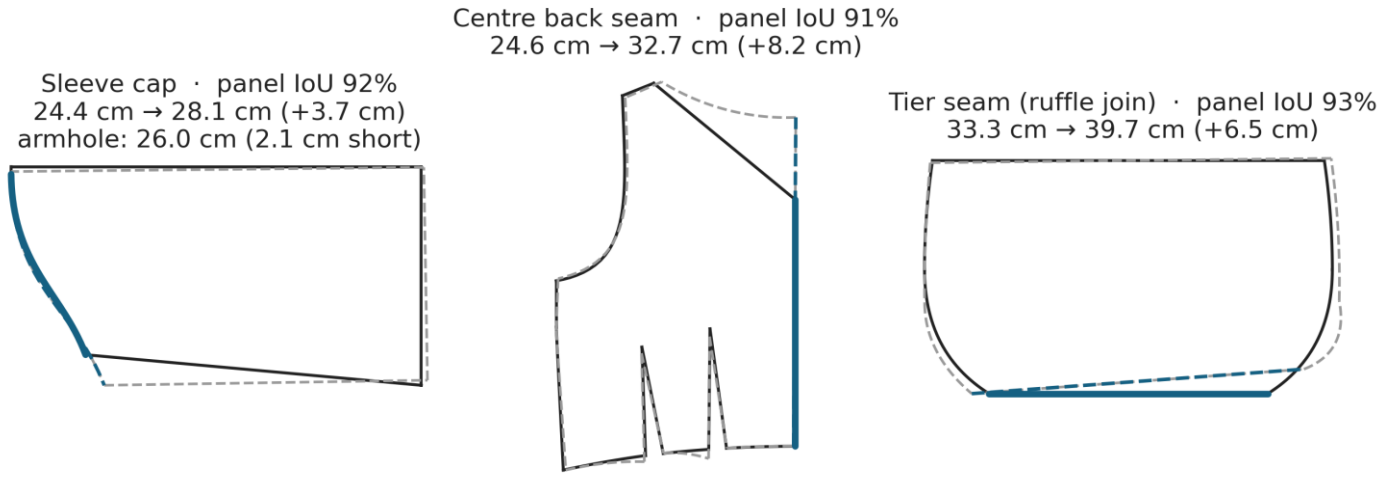
Structure metrics

Metric	Alpparel
Panel names matched	95.7%
Edge slots matched	91.9%
Panel count accuracy	80.4%
Panel IoU	72.0%

Length error by edge role

Role	Mean error	> 3 cm off
tier_seam	254 mm	86%
hemline	65 mm	51%
center_back_seam	62 mm	57%
sleeve_top_seam	52 mm	39%
sleeve_cap ↔ armhole	30 ↔ 16 mm	36 ↔ 16%
dart_leg	13 mm	10%

Only roles with the same number of pieces in ground truth and prediction are compared.



black = ground truth, grey dashed = Alpparel, teal = the edge compared (same panel, role and index)

Panels overlap by 91–93%, yet one edge per panel is 3.7–8.2 cm off. The sleeve cap grows to 28.1 cm while the armhole it is sewn to is 26.0 cm, so the sleeve no longer fits; in the ground truth both are 24.4 cm.

Without roles, these errors average into one number; with roles they are located by edge role.

Known-Error Test: Panel L2 vs Role Length Error

We assume a prediction that is correct except for one edge: one stitched edge of the ground truth is lengthened by δ (the outline stays closed), and both metrics are measured against the original pattern.

Panel L2 (no labels needed)

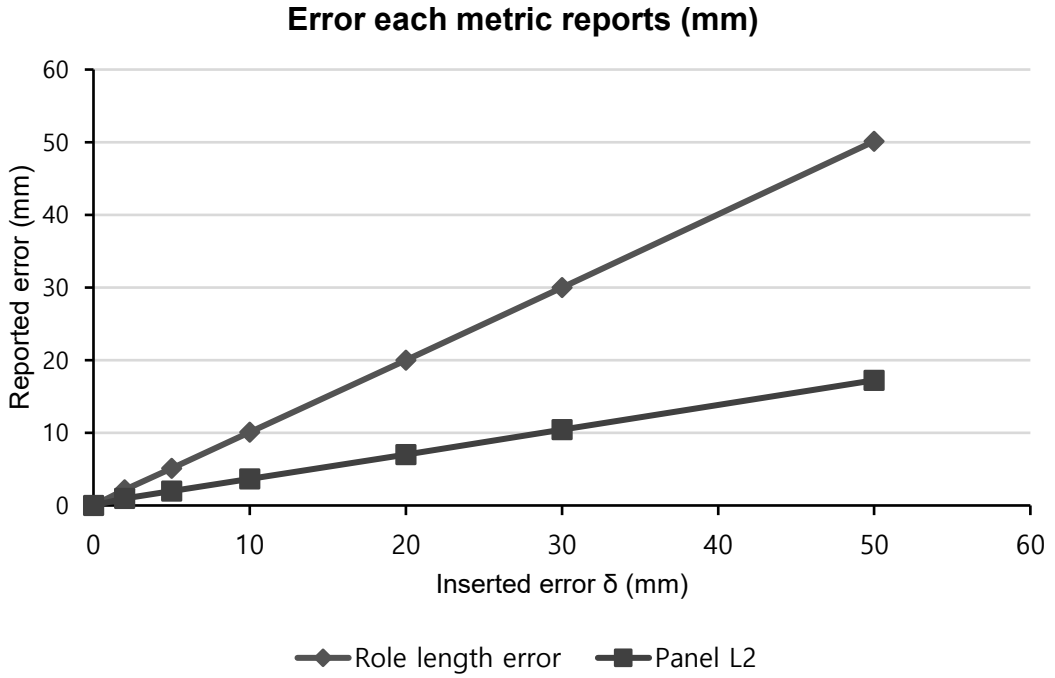
Resample the true and predicted outline of the edited panel into 64 evenly spaced points, align the start, and average the point-to-point distance.

Role length error (needs SewRole)

Pair each predicted edge with the true edge in the same slot (panel, role, index) and take |L_pred − L_true| for the edited edge.

Inserted error δ	Panel L2	Role length error
10 mm	3.6 mm	10.1 mm
30 mm	10.4 mm	30.0 mm
50 mm	17.2 mm	50.1 mm

593 of the 616 benchmark garments (23 have no stitched edge of the tested roles). Edited roles: side 249, waist 210, shoulder 55, sleeve cap 33, armhole 28, underarm 18.



For a 30 mm error, Panel L2 reports 10.4 mm because the error is spread over the whole outline; role length error reports 30.0 mm on the edge that is actually wrong.

Limitations and Future Work

Model coverage

Current role-level benchmark results are shown for AIpparel; broader model comparisons remain to be completed.

LLM application

Further investigation is needed to determine whether geometric metric errors can also be observed in LLM applications like GPT and Claude.

Diagnostic only

The proposed model establishes a new benchmark, which clarifies the problem but does not provide an improvement.

1 Expanded Model Comparison

Evaluate a wider range of state-of-the-art sewing-pattern generation models (e.g., diffusion-based, graph-based) on the proposed benchmark to provide a more comprehensive comparison.

2 GPT and Claude Prompting

Test prompting strategies with leading GPT and Claude models for sewing-pattern generation and edge-level geometry correction.

3 Edge-Level AI Models

Develop AI models that detect and correct geometric problems in individual pattern edges.

References

NeuralTailor
Korosteleva & Lee. NeuralTailor: Reconstructing Sewing Pattern Structures from 3D Point Clouds of Garments. ACM TOG 41(4), 2022.

GarmentCode
Korosteleva & Sorkine-Hornung. GarmentCode: Programming Parametric Sewing Patterns. ACM TOG 42(6), 2023.

GarmentCodeData
Korosteleva et al. GarmentCodeData: A Dataset of 3D Made-to-Measure Garments with Sewing Patterns. ECCV 2024.

Zenodo 3D Garments
Korosteleva & Lee. Generating Datasets of 3D Garments with Sewing Patterns. NeurIPS Datasets and Benchmarks 2021.

SewingLDM
Liu et al. Multimodal Latent Diffusion Model for Complex Sewing Pattern Generation. ICCV 2025.

AIpparel
Nakayama et al. AIpparel: A Multimodal Foundation Model for Digital Garments. CVPR 2025.

Design2GarmentCode
Zhou et al. Design2GarmentCode: Turning Design Concepts to Tangible Garments Through Program Synthesis. CVPR 2025.

Sewformer
L. Liu, X. Xu, Z. Lin, J. Liang & S. Yan. Towards Garment Sewing Pattern Reconstruction from a Single Image. ACM TOG 42(6), 2023 (SIGGRAPH Asia).

ChatGarment
S. Bian et al. ChatGarment: Garment Estimation, Generation and Editing via Large Language Models. CVPR 2025, 2924–2934.

GarmentDiffusion
X. Li, Q. Yao & Y. Wang. GarmentDiffusion: 3D Garment Sewing Pattern Generation with Multimodal Diffusion Transformers. IJCAI 2025.

GarmentGPT
Weng et al. GarmentGPT: Compositional Garment Pattern Generation via Discrete Latent Tokenization. ICLR 2026.

GenPattern
H. Gui, Z. Yang, A. R. Harish & C. Ren. GenPattern: Dual-Graph Enhanced Sewing Pattern Generation via Multimodal LLM. J. Manufacturing Systems 83, 822–838, 2025.

Cao et al.
C. Cao, R. Li, C. Dumery & H. Li. Learning Sewing Patterns via Latent Flow Matching of Implicit Fields. SIGGRAPH 2026.

AutoSew
P. Ríos-Navarro, E. Garces & J. Lopez-Moreno. AutoSew: A Geometric Approach to Stitching Prediction with Graph Neural Networks. WACV 2026.

SwiftTailor
P. Pham, U. D. Tran, B.-S. Hua & P. Nguyen. SwiftTailor: Efficient 3D Garment Generation with Geometry Image Representation. CVPR 2026.

SewFusion
SewFusion: Tailored Generation of Topology and Panel-Level Geometry for Sewing Patterns. arXiv:2609.23548, 2026.

Arc approximation
Dokken, Dæhlen, Lyche & Mørken. Good Approximation of Circles by Curvature-Continuous Bézier Curves. CAGD 7, 1990.

Size intervals
Y. S. Kang. An Analysis of the Distribution and the Body Size of Men According to Shirt Silhouette, Target Age, and Shirt Size. Journal of the Korean Society of Fashion Design 12(4), 81–100, 2012.

Standards
ISO 8559-1:2017 Size designation of clothes — Part 1: Anthropometric definitions for body measurement; ISO 4916:1991 Textiles — Seam types — Classification and terminology; KS K 0050 garment sizing.