

Dressing The Imagination: A Dataset For Ai-powered Translation Of Text Into Fashion Outfits And A Novel NeRA Adapter For Enhanced Feature Adaptation

Gayatri Deshmukh, Somsubhra De, Chirag Sehgal, Jishu
Sen Gupta, Sparsh Mittal

WACV 2026 (Application Track)



CONTENTS

- Motivation
- Contribution
- FLORA Dataset
- NeRA Adapter
- Quantitative Results
- Qualitative Results
- Summary

MOTIVATION

- Fashion design needs fine-grained control (silhouette, fabric, fit, embellishments, accessories).
- General text-to-image models often miss subtle styling details in sketch-style outputs.
- A key bottleneck: lack of curated sketch–text datasets with professional fashion vocabulary.
 - Limited ability of models to adapt non-linearity.

What Paper Contributes

FLORA Dataset

- 4,330 curated sketch–text pairs
- Descriptions use industry jargon (materials, construction, fit)
- Designed for text→fashion sketch generation

NeRA Adapter

- Nonlinear low-rank PEFT using KAN-style basis functions
- Improves fidelity + semantic alignment vs LoRA/LoKR/DoRA/LoHA

FLORA Dataset

Scale

4,330

curated sketch–text pairs

Language

Industry Specific

domain jargon & styling details

- 9 semantic categories cover garment type, materials, detailing, and finishing.
- Descriptions include model pose + outfit details + colors + accessories.

Selected classes from each of the 9 categories in FLORA dataset.



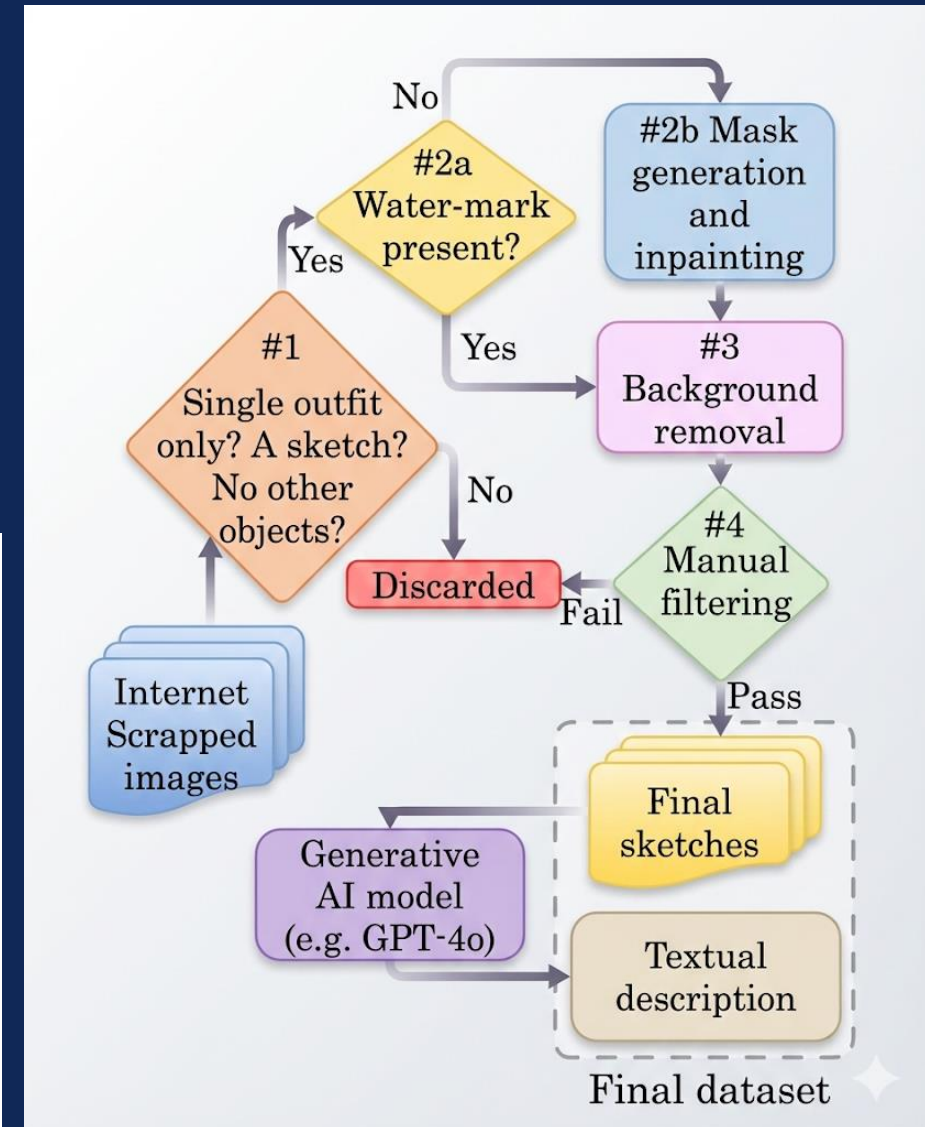
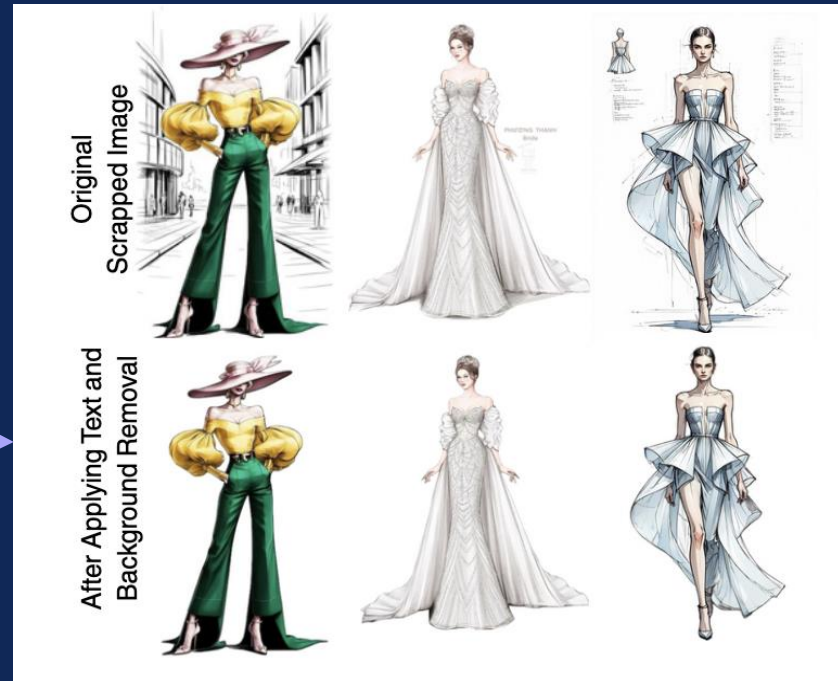
How FLORA was built ?

From web-scraped sketches → clean training data

6

- Web-scrape: 10,042 candidate sketch images
- Prompt filter (VLM use): filter single outfit + sketch-only images
- Watermark removal: use of OCR mask + diffusion inpainting
- Background cleanup to remove unwanted objects (potential noise)
- manual deduplication removal
- Generate rich descriptions (GPT-4o) → final pairs

Visual comparison of original and processed images.



Sample Image-Text Pair from FLORA



Model Presence: Present

Model Gender: Female

Model Pose: The model is standing with one leg slightly forward. The head is held upright, facing forward.

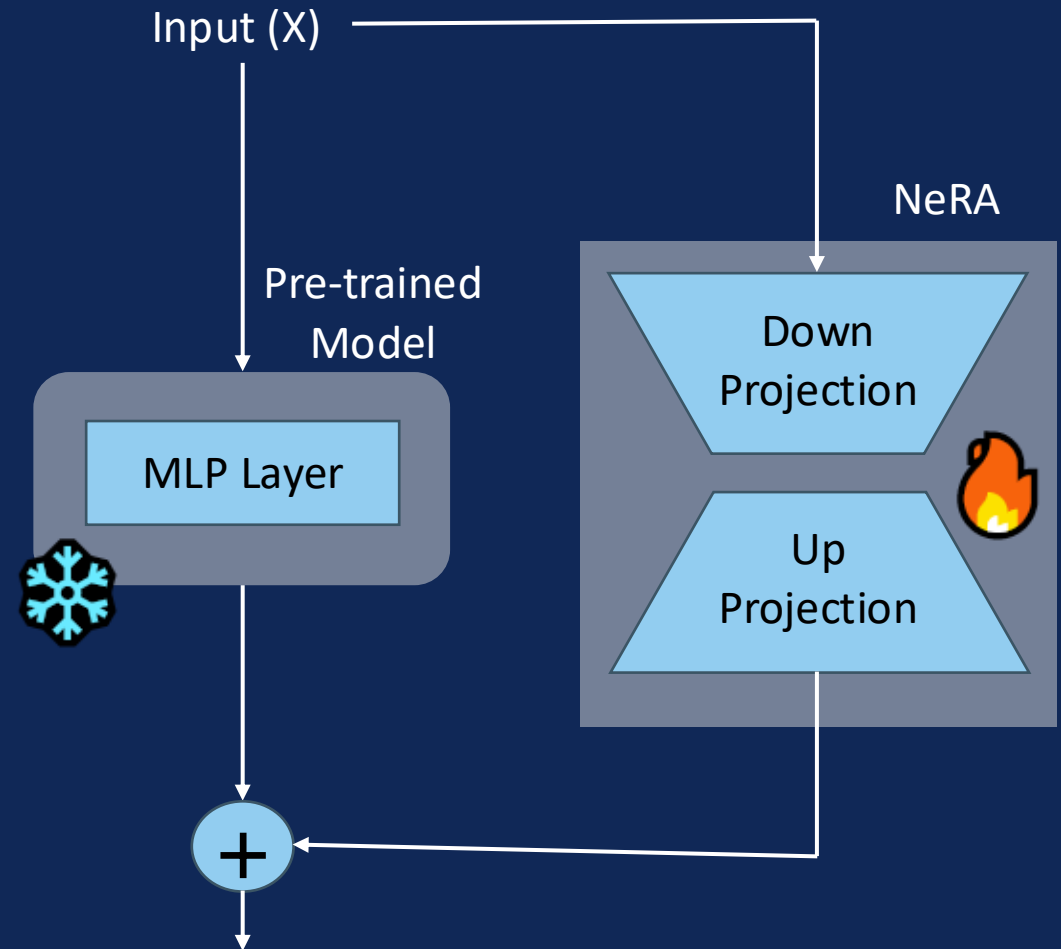
Outfit Description: The outfit is a floor-length evening gown crafted from a luxurious silk satin, known for its lustrous finish and fluid drape. The gown features a sweetheart neckline, complemented by delicate, adjustable spaghetti straps that enhance the silhouette. The bodice is form-fitting, designed to accentuate the waist, while the skirt flows into a soft A-line silhouette, allowing for graceful movement. The gown is tailored to fit snugly at the bust and waist, transitioning into a more relaxed fit through the hips, ensuring a flattering yet comfortable wear. The overall silhouette is structured yet feminine, embodying elegance and sophistication, perfect for formal occasions.

Color Details: The outfit features a vibrant primary color of coral orange, complemented by subtle highlights of lighter peach tones. The fabric exhibits a smooth gradient effect, transitioning from deeper orange at the bodice to lighter shades towards the hem, with a shimmering embellishment at the waist adding a touch of sparkle.

Accessories: No additional accessories present.

NeRA: Nonlinear Low-rank Expressive Representation Adapter

- Standard PEFT Limitations
 - Current adapters (LoRA, DoRA) use MLPs (Multi-Layer Perceptrons).
 - Issue: MLPs are often too linear/shallow to capture the finegrained, non-linear details. (Like one present in fashion text e.g. lace texture, complex folds names etc)
- NeRA uses Uses learnable spline-based activations (B-splines/RBFs) instead of fixed activation functions.
- This allows the model to "bend" more to fit complex data patterns.



Fixed activations can't bend enough, fully learnable activations can bend too far - NeRA keeps the MLP fixed and learns trainable spline activations only in the adapter to get controlled, high-power adaptation.



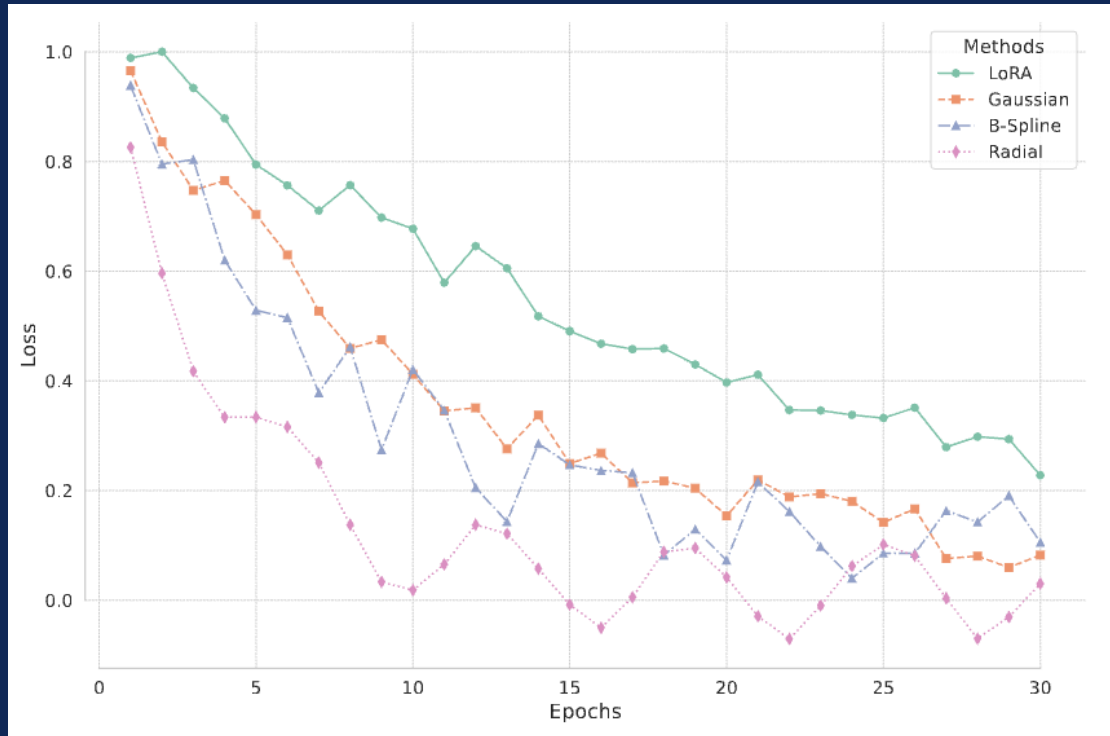
Quantitative Results

Methods	Zero Shot FID(↓)	Zero Shot CLIPSIM(↑)	LoHA FID(↓)	LoHA CLIPSIM(↑)	DoRA FID(↓)	DoRA CLIPSIM(↑)	LoKR FID(↓)	LoKR CLIPSIM(↑)	LoRA FID(↓)	LoRA CLIPSIM(↑)	NeRA FID(↓)	NeRA CLIPSIM(↑)
FLUX	09.09	0.3112	09.34	0.3001	09.01	0.3131	08.33	0.3223	07.88	0.2987	06.05	0.3412
SD	19.25	0.0623	20.01	0.0630	19.02	0.0720	19.33	0.0751	17.32	0.0910	16.51	0.1112
SD-XL	16.37	0.1539	14.47	0.1667	15.89	0.1523	14.02	0.1782	15.32	0.1522	13.33	0.1897
SD-3	13.28	0.2118	13.62	0.2337	13.19	0.2279	13.01	0.2555	09.37	0.2119	08.07	0.2503
Pixart-Sigma	14.11	0.2666	13.88	0.2424	13.88	0.2707	12.92	0.2992	09.89	0.2542	08.28	0.2991

LoHA, DoRA, LoKR, LoRA and NeRA are finetuned on **FLORA** dataset.

Efficiency and Convergence

Faster fine-tuning with modest overhead



Model	Latency (s)↓		Parameters(M)↓		Training Time (Hrs)↓	
	LoRA	NeRA	LoRA	NeRA	LoRA	NeRA
FLUX	8.73	8.81	2.97	3.43	14.57	6.78
SD	13.43	14.42	1.47	1.82	9.37	5.62
SD-XL	9.87	10.02	2.81	3.65	9.23	4.97
SD-3	10.11	10.23	3.11	3.92	11.47	4.12
Pixart	6.32	6.89	0.89	1.12	12.47	6.01

Qualitative Results...

Input Prompt

Create a fashion outfit sketch based on the following specifications:

MODEL PRESENCE: Present

MODEL GENDER: Female

MODEL POSE: Standing tall with **one hand resting on her hip**, the other gracefully extended to the side. Her **head is tilted slightly back**, exuding elegance and confidence. Overall stance is poised and statuesque.

OUTFIT DESCRIPTION: A **floor-length ball gown** made from **velvety satin**, featuring a plunging **V-neckline with delicate lace overlays**. The bodice has structured boning, cinching at the waist and flowing into a voluminous skirt with **cascading layers of tulle**. The **back features a dramatic open design with crisscross straps**.

COLOR DETAILS: The gown is **royal blue** with **silver embroidered accents**, creating a shimmering effect under light.

ACCESSORIES: The model wears a **diamond choker**, matching **stud earrings**, and **silver stiletto heels**. Her hair is styled in soft, **romantic curls**. No additional accessories are present.









































Zero Shot Output (Without Finetuning on FLORA)



NeRA Output (After Finetuning on FLORA)



Qualitative Results

FLUX					Stable Diffusion 3				
Zero Shot	DoRA Adapter	LoKR Adapter	LoRA Adapter	NeRA Adapter	Zero Shot	DoRA Adapter	LoKR Adapter	LoRA Adapter	NeRA Adapter
<p>(1) Silhouette: floor-length ball gown made from velvety satin, V-neckline with delicate lace overlays. Color Palette: royal blue with silver embroidered accents, Shimmering effects. Accessories: diamond choker.</p>									
									
<p>(2) Silhouette: fitted blazer over a silk blouse, paired with high-waisted pencil skirt that hits just below the knee. Color Palette: Blazer and skirt are midnight navy, blouse is ivory white. Accessories: ear stud earrings, classic wristwatch.</p>									
									
<p>(3) Silhouette: flowy sundress with thin spaghetti straps and sweetheart neckline, skirt has tiered ruffles. Color Palette: pastel yellow with floral pink and lavender prints. Accessories: Straw hat, small hoop earrings, and strappy sandals.</p>									
									
<p>(4) Silhouette: mermaid gown, one-shoulder neckline, asymmetrical ruffles cascading over the shoulder. Color Palette: emerald green base. Accessories: earrings, a single silver cocktail ring, and matching metallic heels.</p>									
									

CONCLUSION

Bridging the Gap

Addressed the “*Design Gap*” in Fashion AI by moving beyond photorealism to conceptual sketching

Built the first robust pipeline to translate industry-specific jargon into high-fidelity visual designs

Core Contributions

FLORA Dataset: Curated 4,330 sketch–text pairs, enabling models to learn nuanced design elements (*e.g., fabric textures, cuts*)

NeRA Adapter: Novel architecture using Kolmogorov–Arnold Networks (KANs); replaces static MLPs with learnable splines, delivering stronger non-linear adaptability than LoRA

THANK YOU.....!

Scan to access code
and dataset.

