

# SVGformer

## Representation Learning for Continuous Vector Graphics using Transformers

Defu Cao<sup>1</sup>, Zhaowen Wang<sup>2</sup>, Jose Echevarria<sup>2</sup>, Yan Liu<sup>1</sup>

<sup>1</sup>University of Southern California

<sup>2</sup>Adobe Research

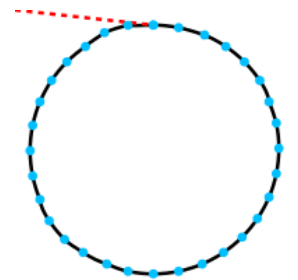
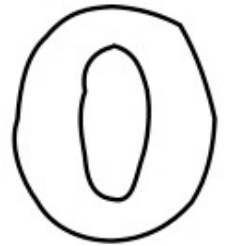


# Overview

SVGformer captures both geometric information and curve semantic information to handle long-term sequence relationships in SVG.

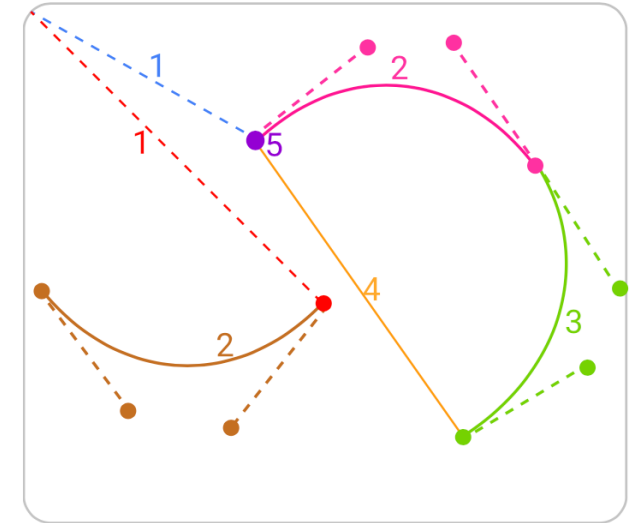
SVGformer can take original continuous format as input which can effectively reduce the token space in the embedding layer.

SVGformer achieves new state-of-the-art performance on four downstream tasks of both Font and Icon dataset.



# Vector Graphics: Representation

- Bitmap
  - Pixel level
  - Hard to represent detail attributes
- Scalable Vector Graphics (SVG):
  - Geometric primitives
    - Lines, circles, polygons, splines
  - Commands of varying length
    - Long term dependencies
  - Fonts
    - Stylistic features per family



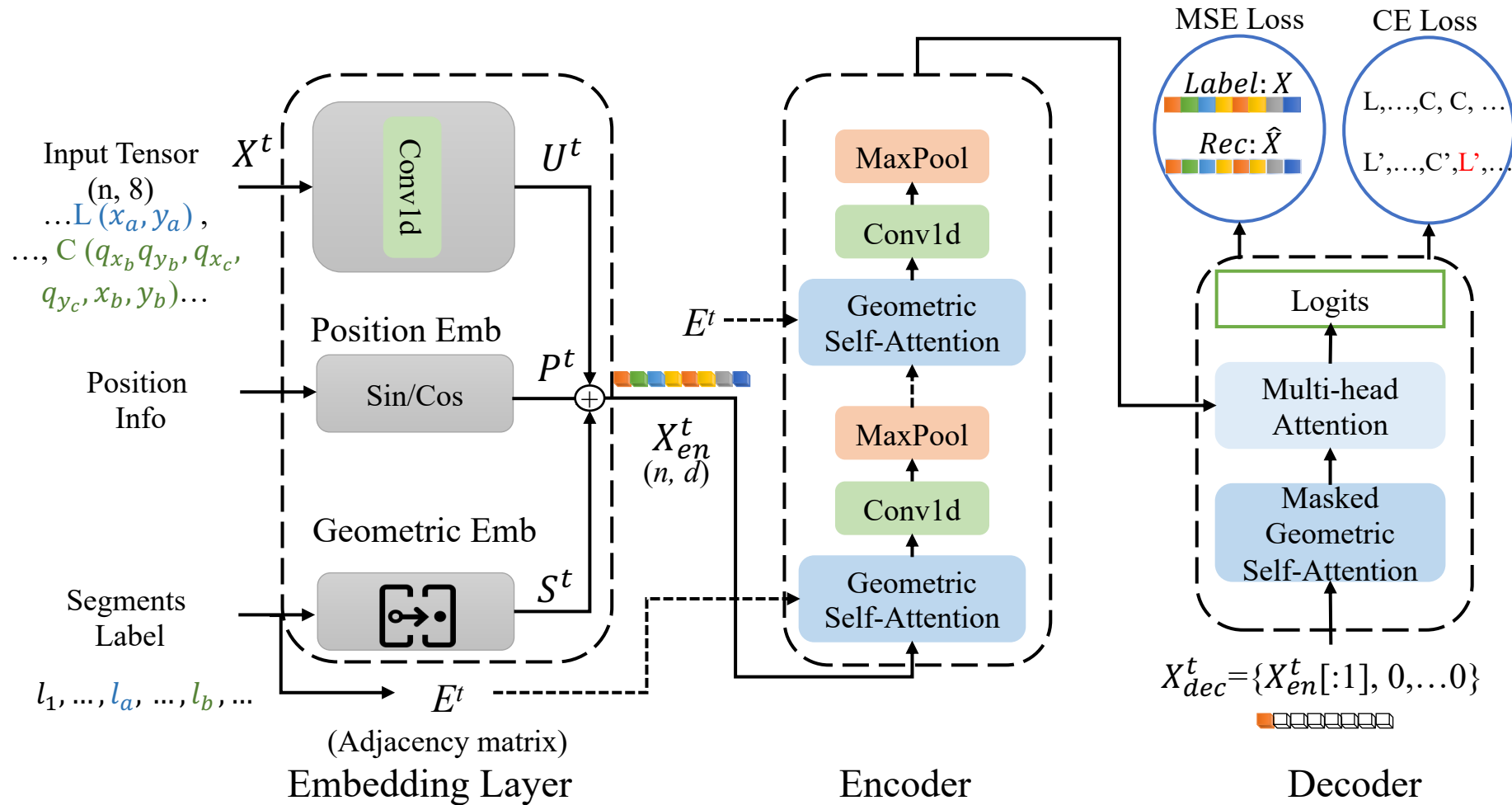
Index	Command	Arguments	$P_1$
0	<SOS>	-1 -1 -1 -1 -1 -1	
1	m	-1 -1 -1 -1 95 54	
2	c	143 16 180 14 215 66	
3	c	255 118 244 152 191 181	
4	l	-1 -1 -1 -1 95 54	
5	z	-1 -1 -1 -1 -1 -1	
6	<EOS>	-1 -1 -1 -1 -1 -1	

# Challenges

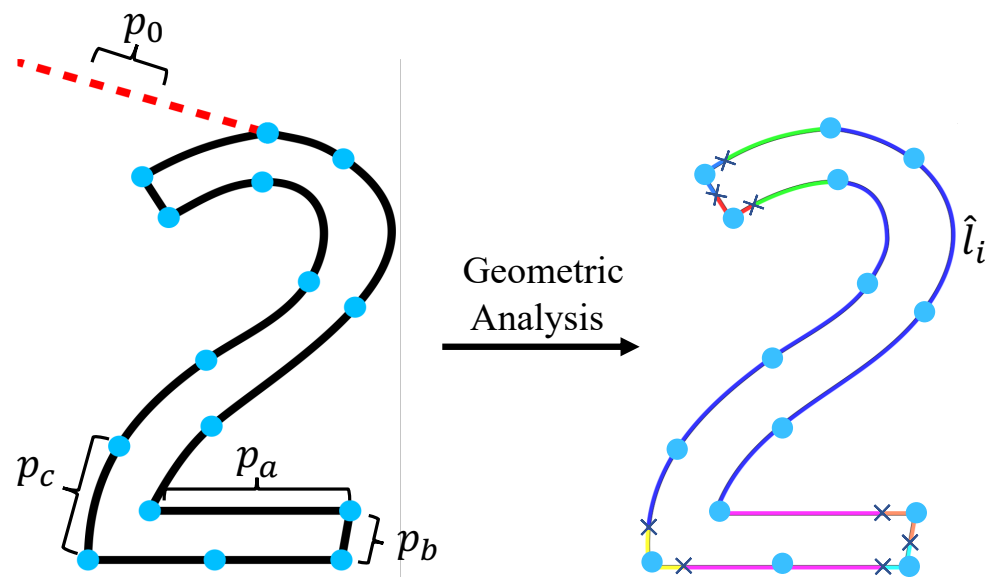
- Learnable Vector tokenization with continuous variables
- Long term sequence relationship:
  - Geometric information for 2D sequence relationship is needed

Methods	Font-MF	SVG-VAE	Im2Vec	DeepVecFont	DeepSVG	DeepLayout	Ours
Encoding Modality	Seq	Img	Img	Img&Seq	Seq	Seq	Seq
Decoding Modality	Seq	Img&Seq	Seq	Img&Seq	Seq	Seq	Seq
Model Arch	GP-LVM	VAE-based	RNN Encoder Decoder	LSTM+CNN Encoder Decoder	Transformer-based	Transformer-based	Transformer-based
Seq Format	Keypoints	Commands	Keypoints	Commands	Commands	Commands	Commands
Geometric information	-	-	-	-	-	-	Segment

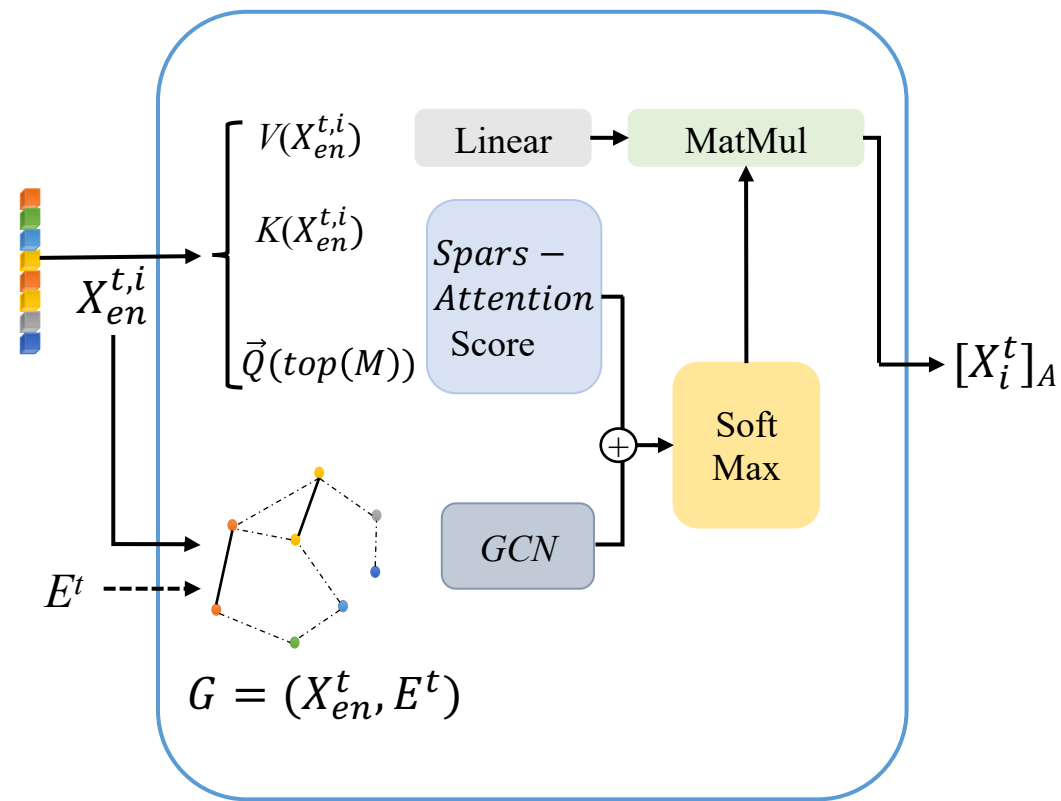
# Method – SVGformer Architecture



# Method – Key Technologies



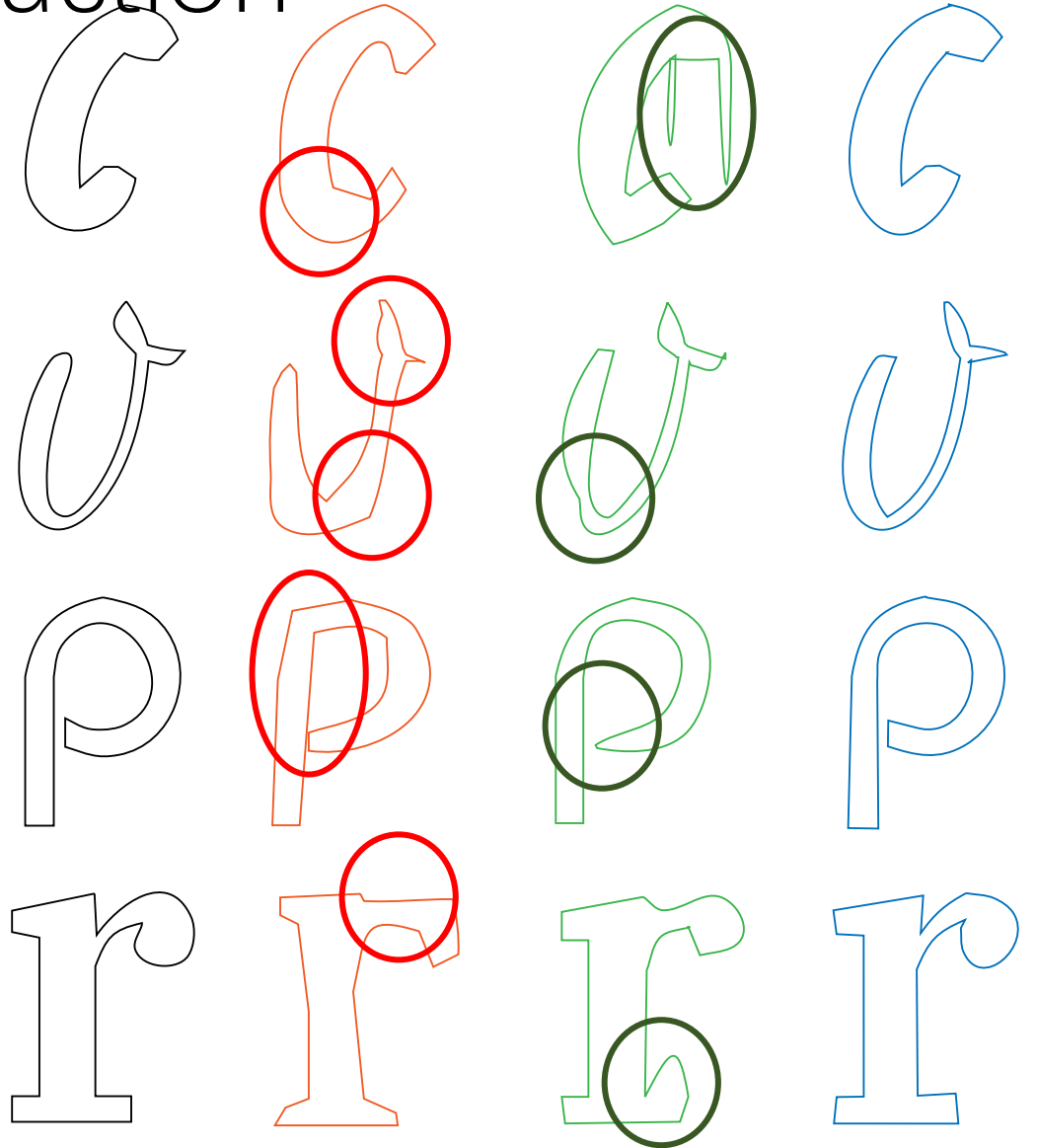
Geometric  
Information Extraction



Geometric  
Self-Attention Module

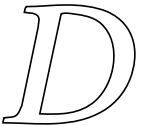






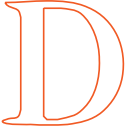




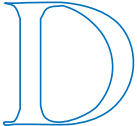








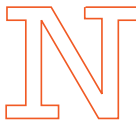








# Results – Fonts, Reconstruction

Reconstruction results on various glyphs.  
From left to right, the original SVG (black),  
DeepSVG (orange), LayoutTransformer  
(green) and our SVGformer (blue),  
respectively.



# Results – Fonts, Retrieval






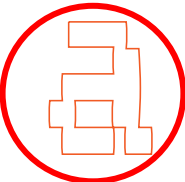






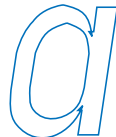


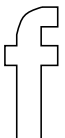














- Retrieval results from all glyphs. Results for DeepSVG are shown in orange, and results for SVGformer in blue.
- SVGformer always retrieves the same glyph as the query when searching from all glyphs.

Query	Top7 retrieval results on all candidates						
							
							
							
							



# Results – Fonts, Retrieval

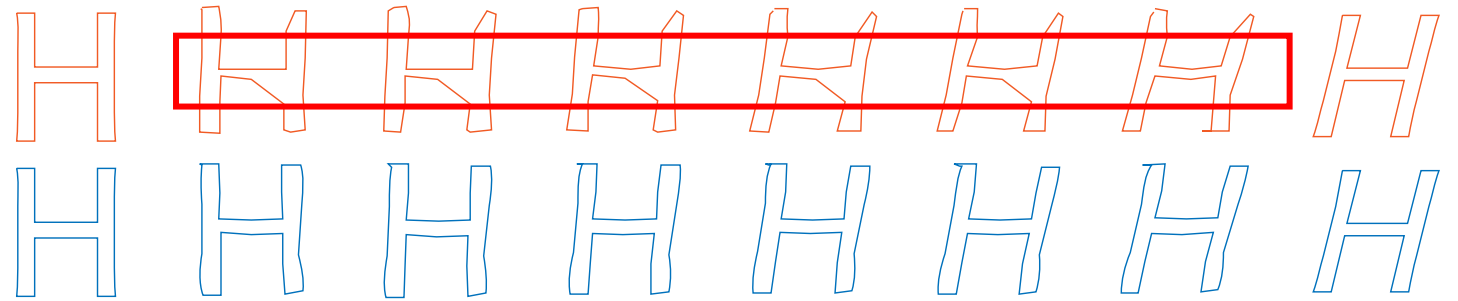
- Retrieval results from all glyphs. Results for DeepSVG are shown in orange, and results for SVGformer in blue.
- SVGformer can return results in more similar styles when searching for a specific glyph.

Query	Top7 retrieval results on same scope						
							
							
							
							

# Results – Fonts, Interpolate

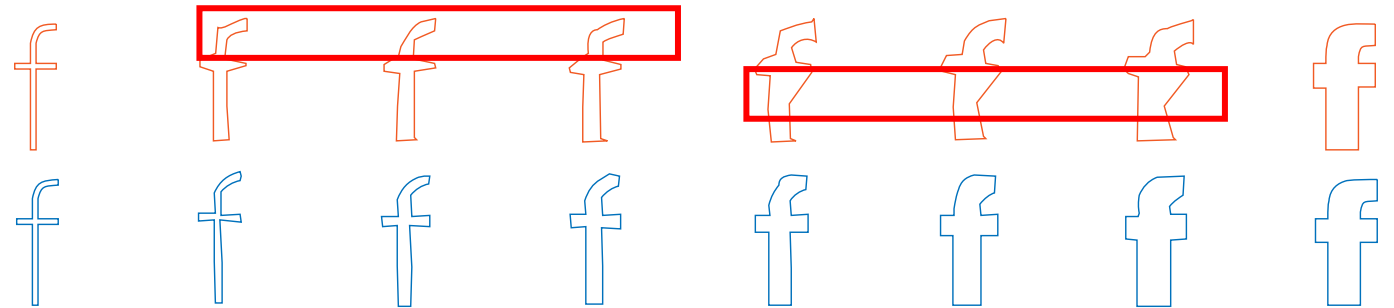
- Interpolations between glyphs. Results for DeepSVG are shown in **orange**, and results for SVGformer in **blue**. The first and last columns are the given glyphs to be interpolated, and the middle columns are the interpolations.

Regular to Italic. *Andika New Basic*, ‘H’



- SVGformer can give more meaningful and smooth results with fewer artifacts and contortions when handling both translations and deformations

Thin to Black. *Saira Semi Condensed*, ‘f’

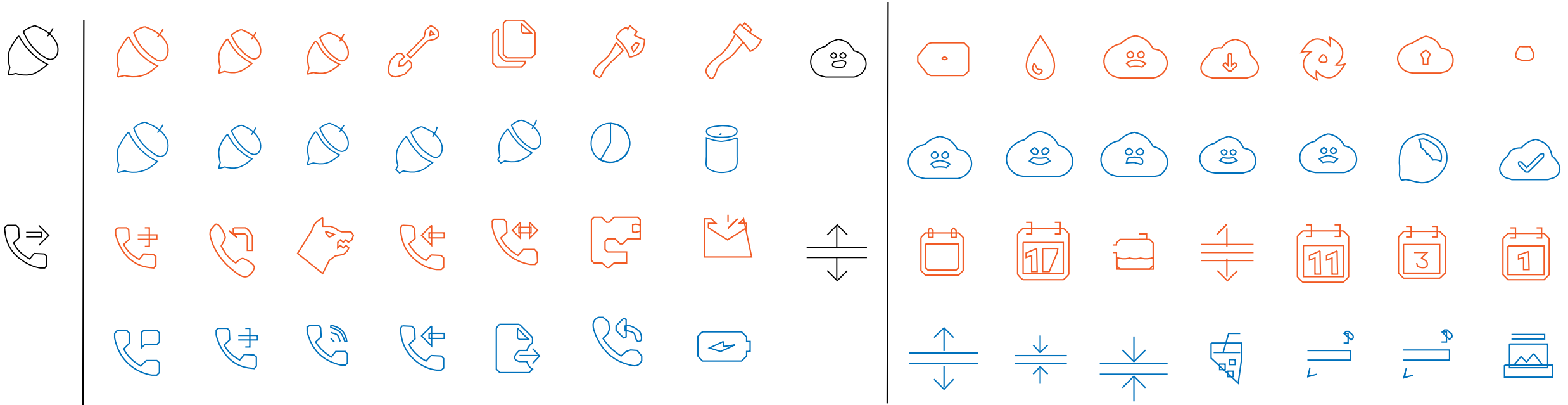


# Results – Icons, Reconstruction

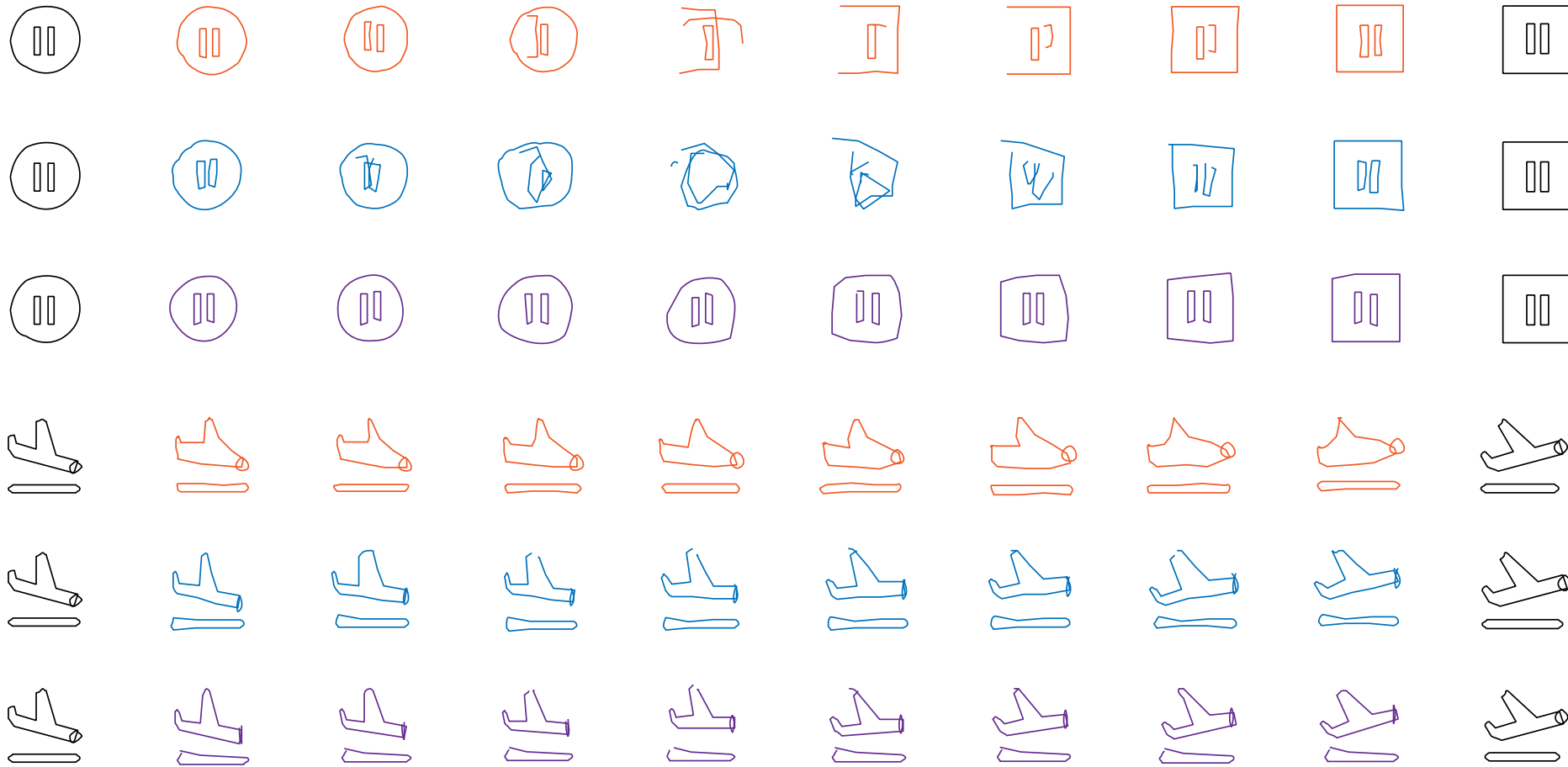
Reconstruction results on Icons. From left to right, the original SVG (black), DeepSVG (orange), LayoutTransformer (green) and our SVGformer (blue), respectively.



# Results – Icons, Retrieval



# Results – Icons, Interpolate



# SVGformer

Representation Learning for Continuous Vector Graphics  
using Transformers

Defu Cao<sup>1</sup>, Zhaowen Wang<sup>2</sup>, Jose Echevarria<sup>2</sup>, Yan Liu<sup>1</sup>

<sup>1</sup>University of Southern California

<sup>2</sup>Adobe Research

