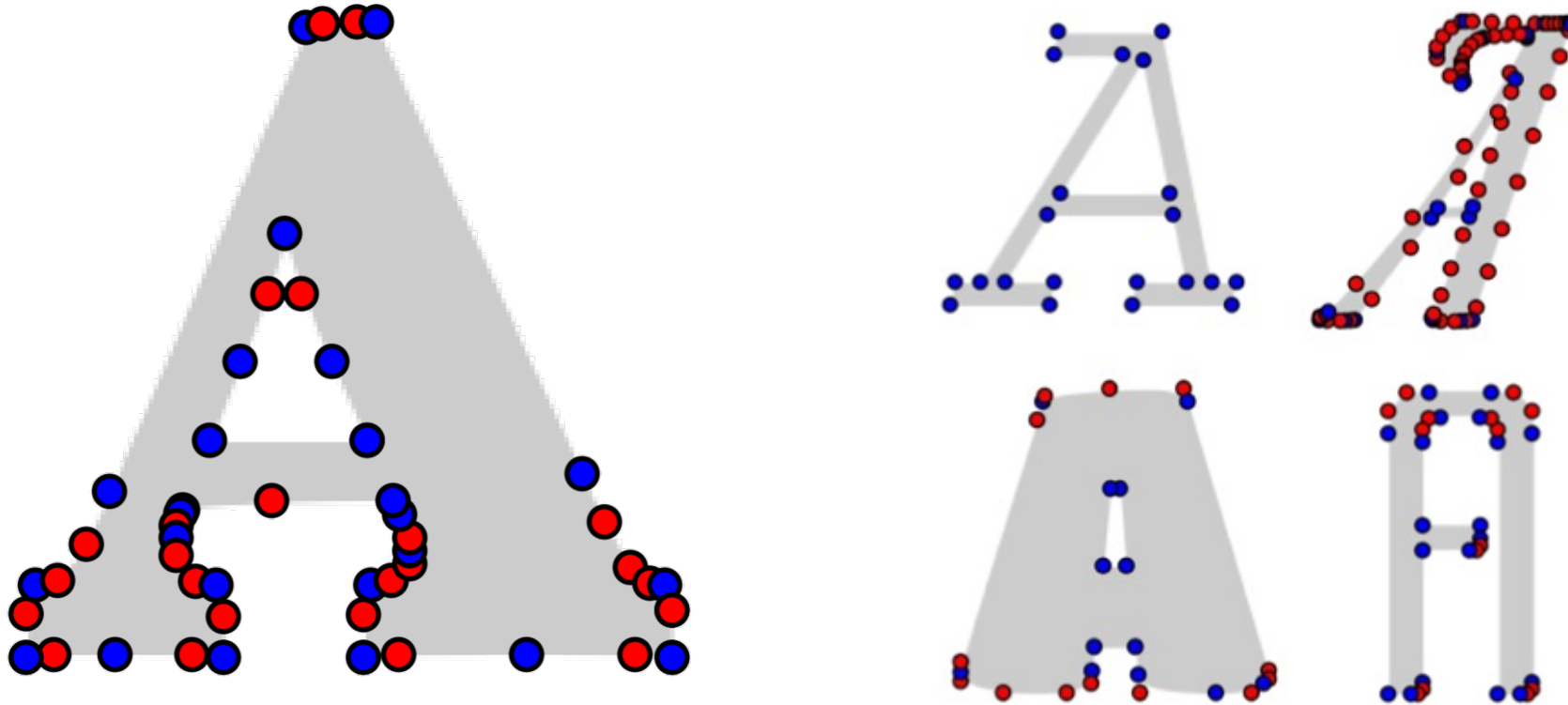


TrueType Transformer: Character and Font Style Recognition in Outline Format

Yusuke Nagata, Jinki Otao,
Daichi Haraguchi, Seiichi Uchida
Kyushu University, Japan

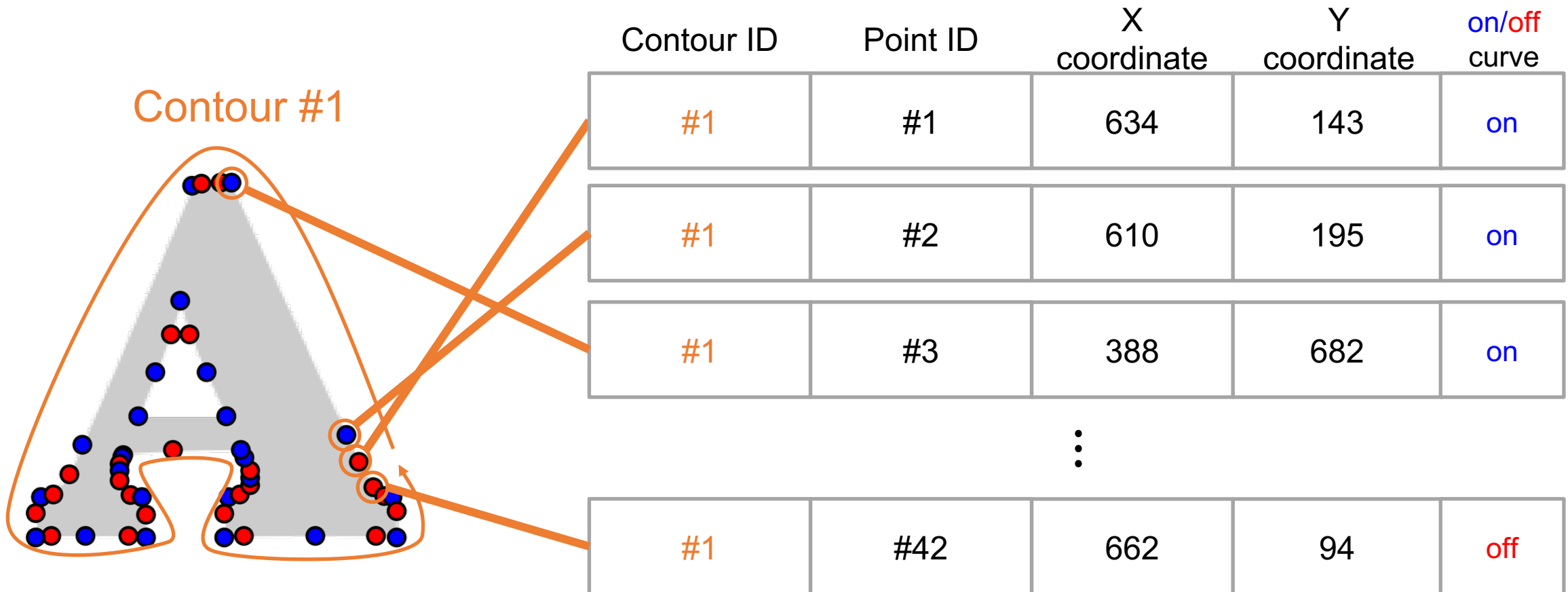
Introduction

Outline is formed by control points



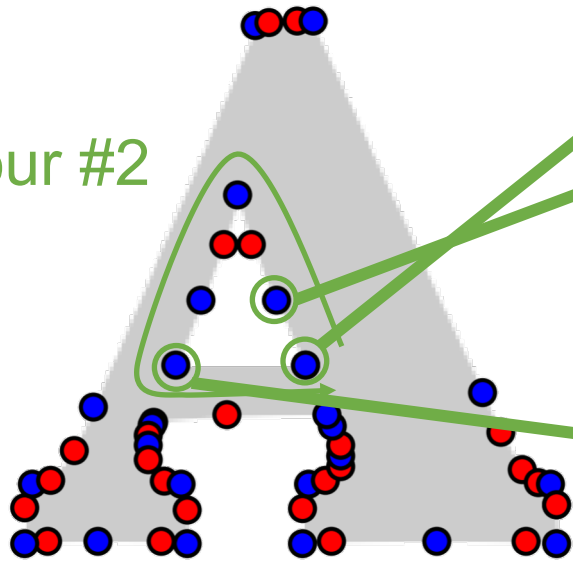
- On-curve point
 - Off-curve point
- } Explained later

Each contour is a sequence of control points



Each contour is a sequence of control points

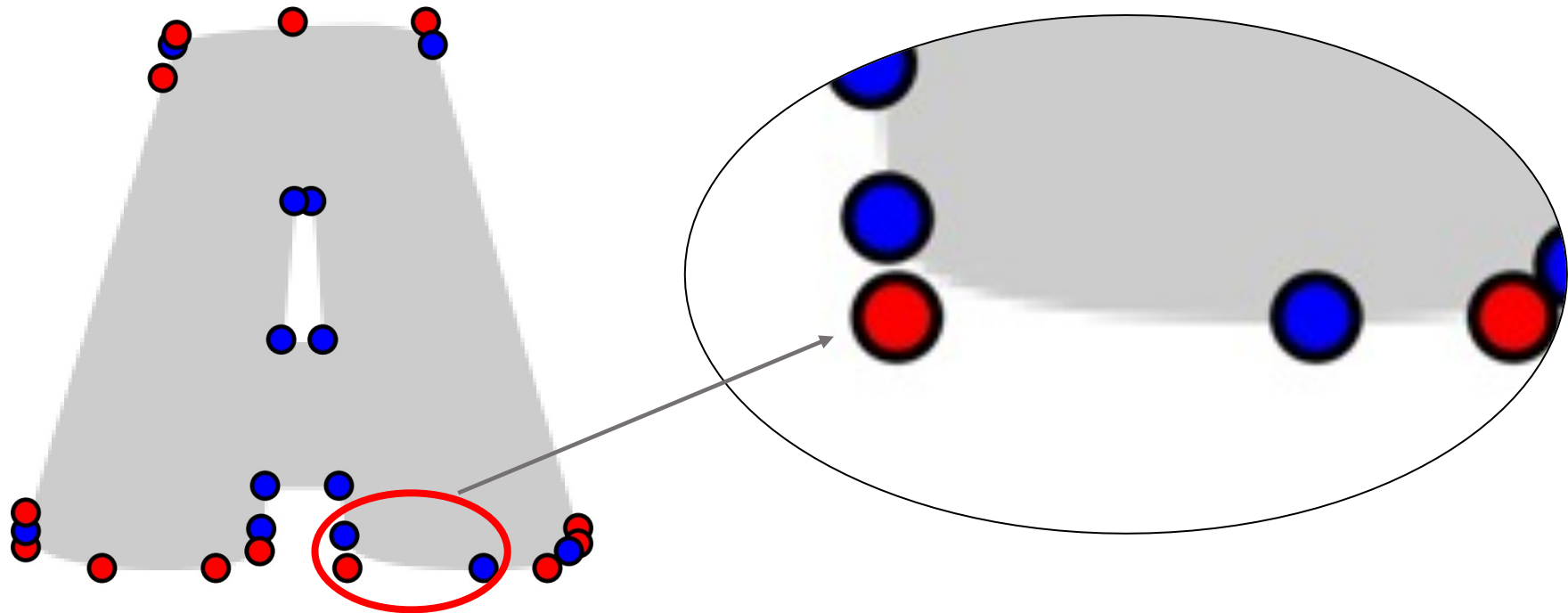
Contour #2



Contour ID	Point ID	X coordinate	Y coordinate	on/off curve
#2	#1	378	231	on
#2	#2	340	316	on
⋮				
#2	#7	208	231	on

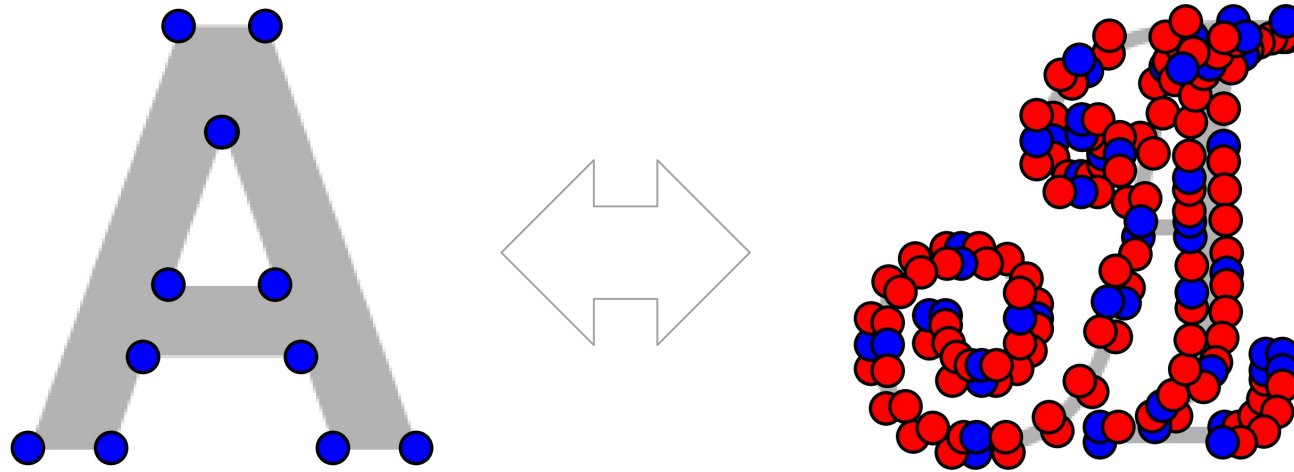
Two types of control points

- On-curve point = Passing points and corners
- Off-curve point = Curvature control point

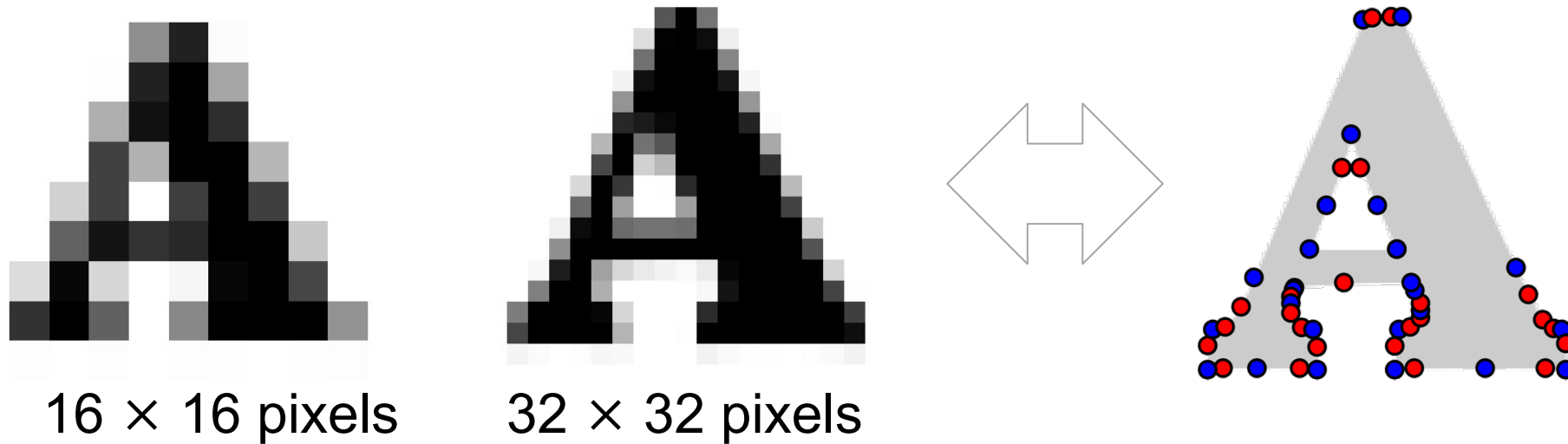


Combination of these types realizes fine and various contour shapes

#Control points is (very) variable

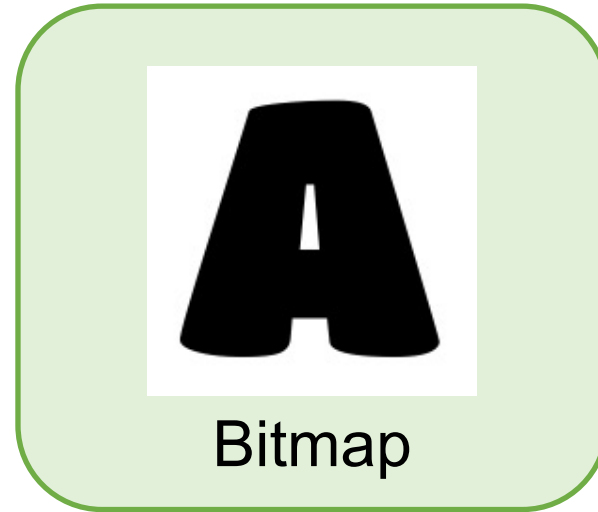
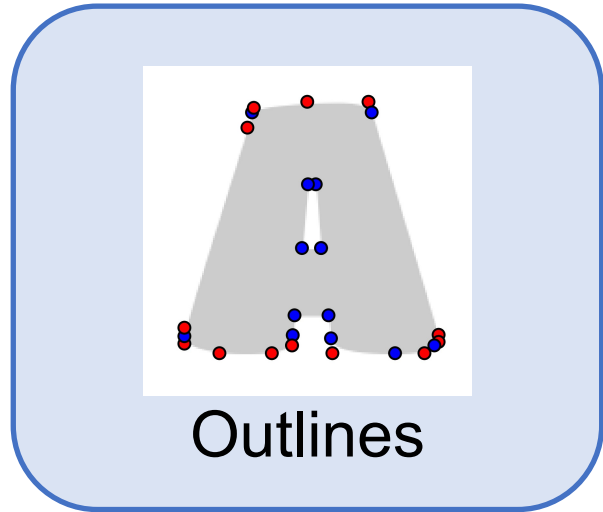


Outline is resolution-free



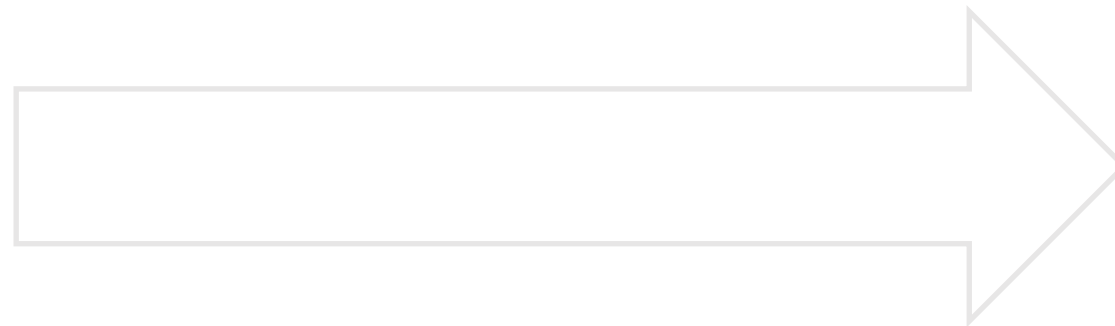
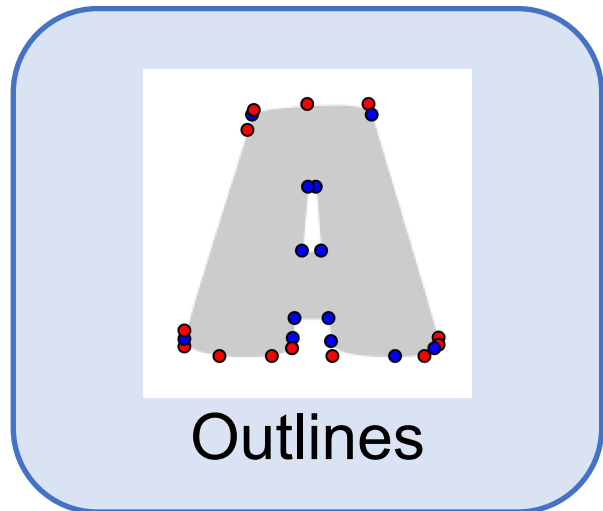
Purpose: Character recognition with outline information

Conventional



Classifier

Ours



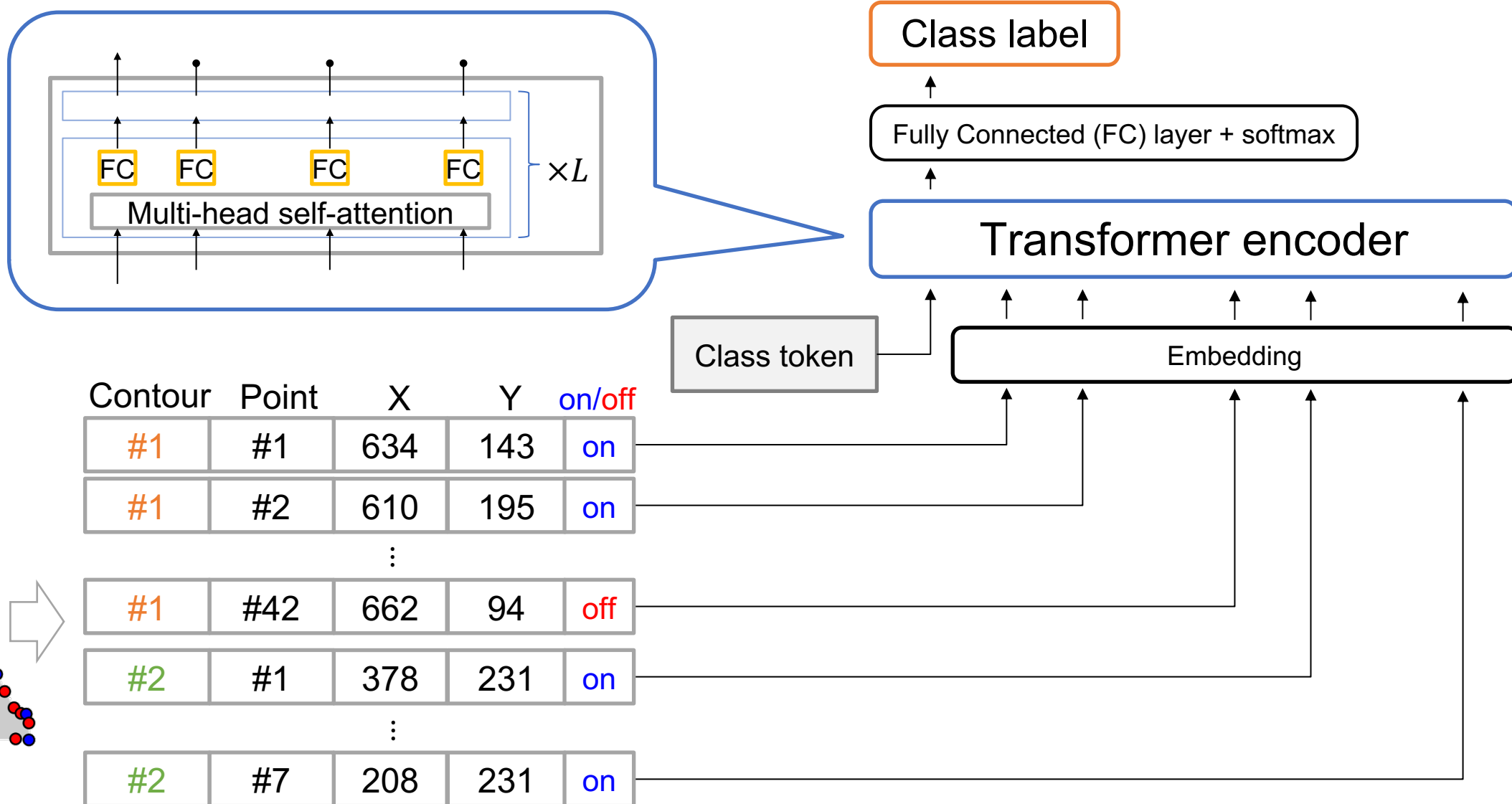
Classifier

Why, how, and when

- Why (merits)
 - **Resolution-free** recognition
 - **Fine shape structure** (good for style recognition)
- How (methodology)
 - **Transformer**, for dealing with variable-length sequences
- When (applications)
 - Direct character recognition of **born-digital documents**
 - Font **style** recognition

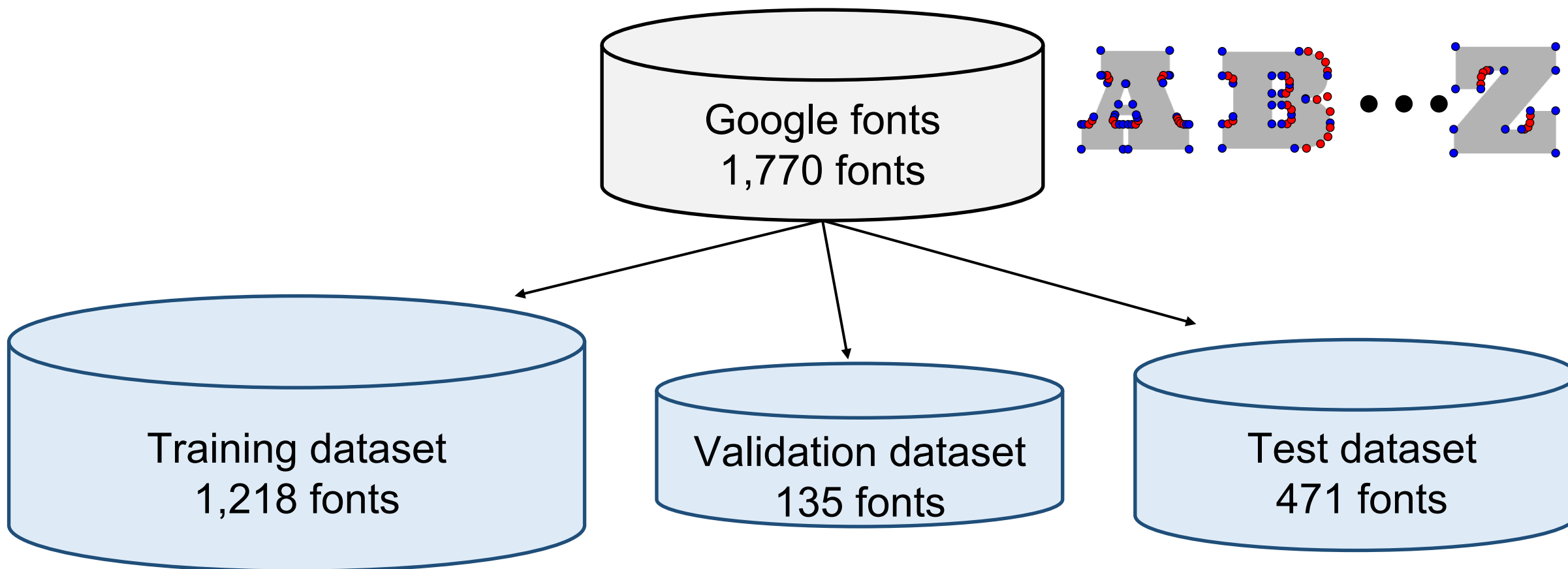
Proposed method:
TrueType Transformer (T^3)

Proposed method: True Type Transformer (T³)



Experiments

Google fonts[*] (True Type Font format)

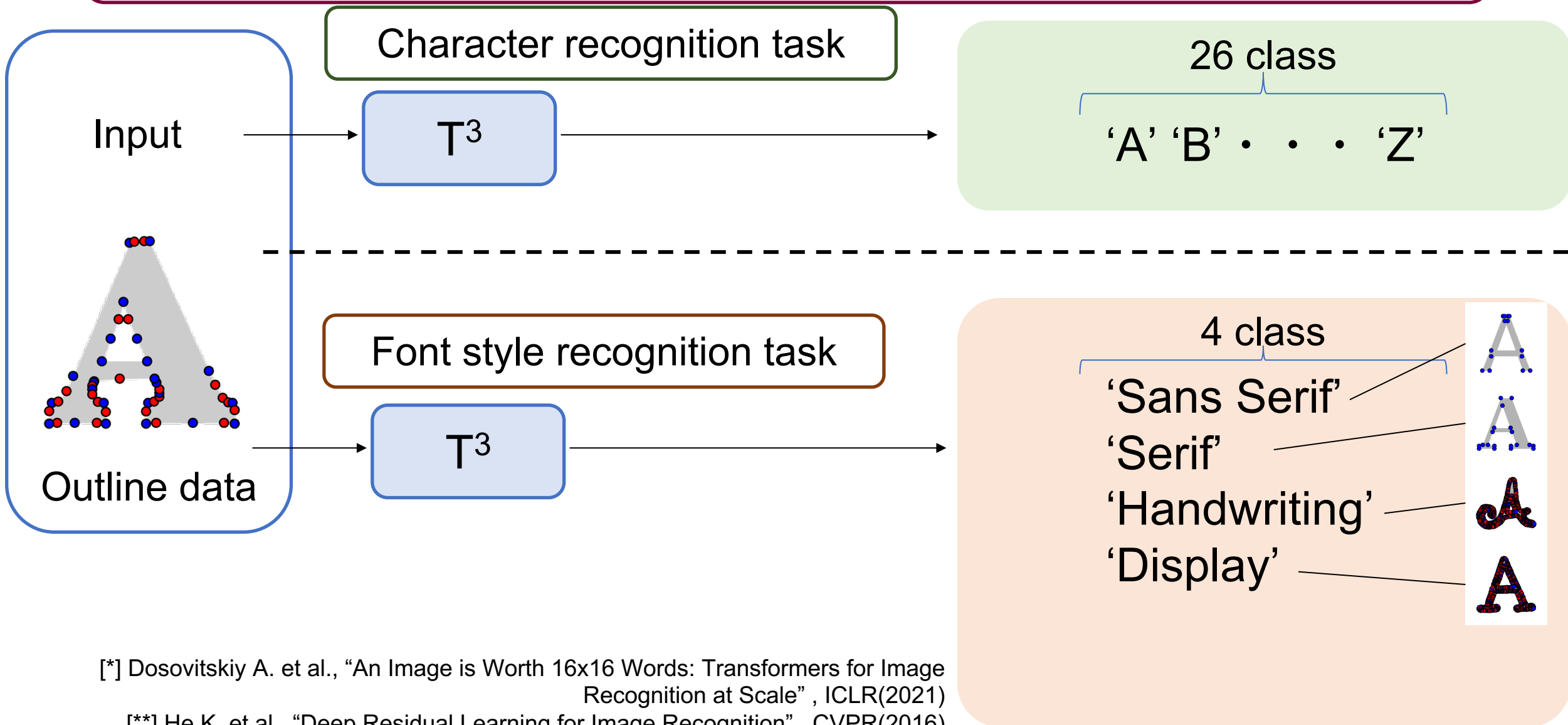


- Followed STEFANN[**] for data split.

[*]<https://github.com/google/fonts>

[**] Roy P. et al., "STEFANN: Scene Text Editor using Font Adaptive Neural Network", CVPR(2020)

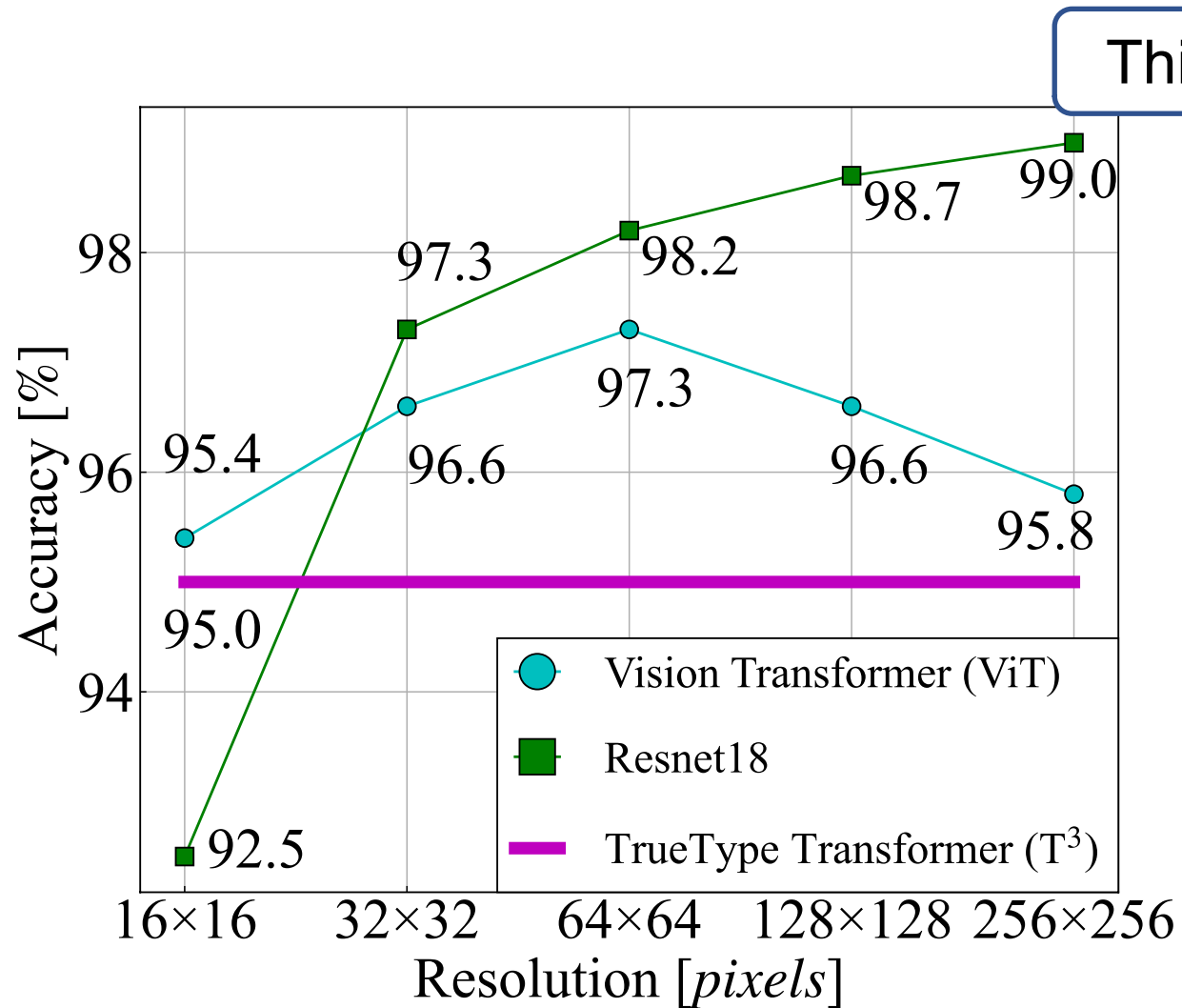
Character & Font style recognition



[*] Dosovitskiy A. et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale", ICLR(2021)

[**] He K. et al., "Deep Residual Learning for Image Recognition", CVPR(2016)

Character recognition accuracy



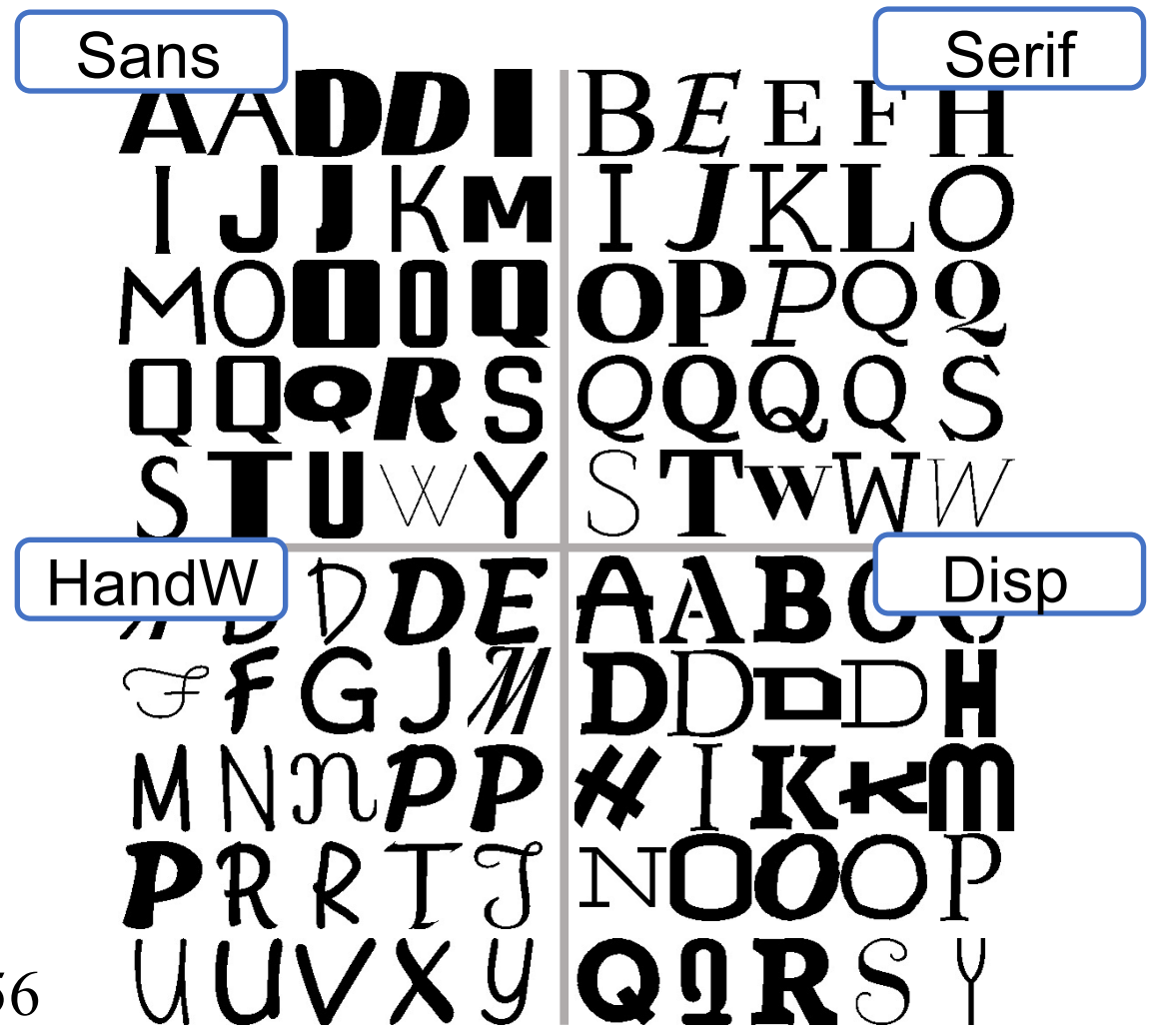
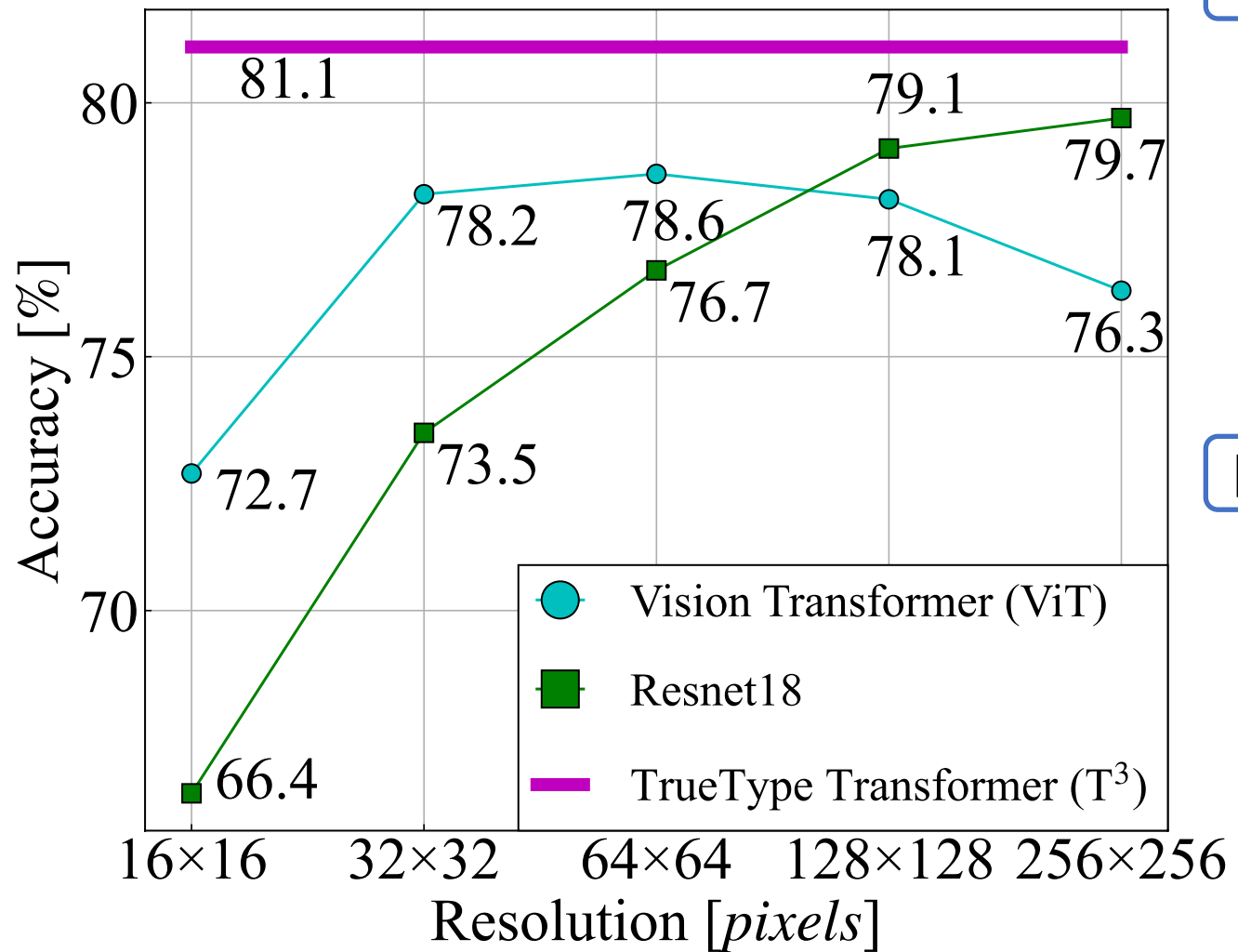
Thin lines



T³: Correct,

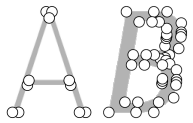
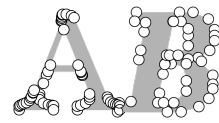

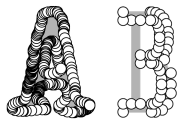



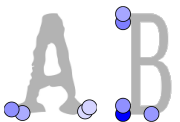
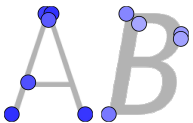


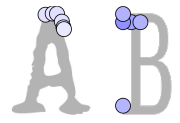
ViT & ResNet18 (64x64): Misrecognized

Font style recognition accuracy



T³: Correct,

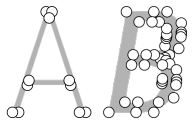


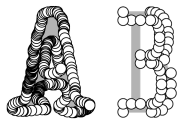
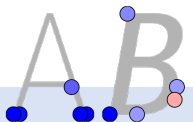
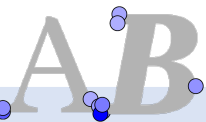

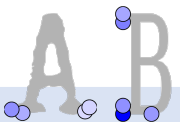
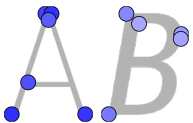


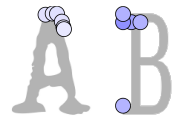
Control points with larger attentions

	Sans serif	Serif	HandW	Disp
All points				
Character recognition				
Font style recognition				

Attention regardless of the styles

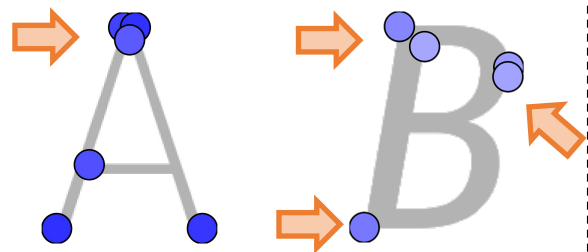
	Sans serif	Serif	HandW	Disp
All points				
Character recognition				
Font style recognition				

Character recognition focuses on the bottom

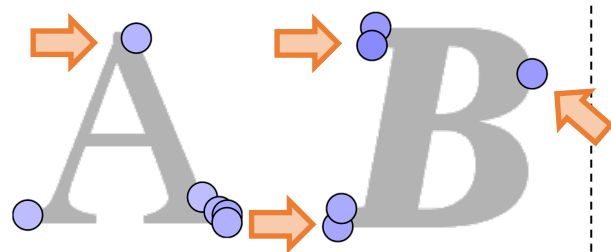
	Sans serif	Serif	HandW	Disp
All points				
Character recognition				
Font style recognition				

Points important for styles

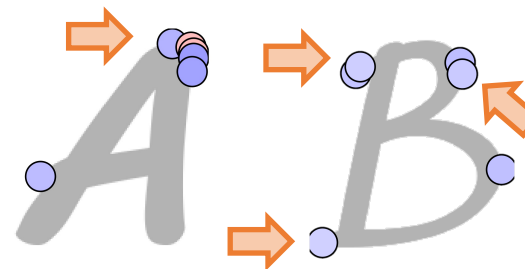
Sans serif



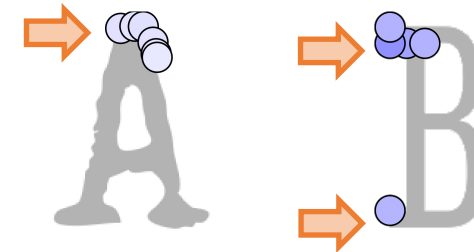
Serif



HandW



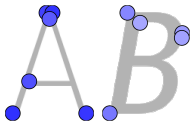


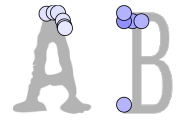
Disp



Conclusion

T^3 can accept a sequence of an arbitrary number of control points

- For the outline-based recognition, we developed T^3 which is a Transformer-based classification model.
- The outline format is suitable for font style recognition.

	Sans serif	Serif	HandW	Disp
Font style recognition				

T^3 can use the fine structure of the font shape.