Sketch Composer

Jiahao Liu Yun Li Zheyuan Zhou

BROWN

GITHUB REPO

INTRODUCTION

Sketch Generating Problem

DATASET

- QuickDraw dataset
- 70K training samples
- 2.5K validation and 2.5K test samples

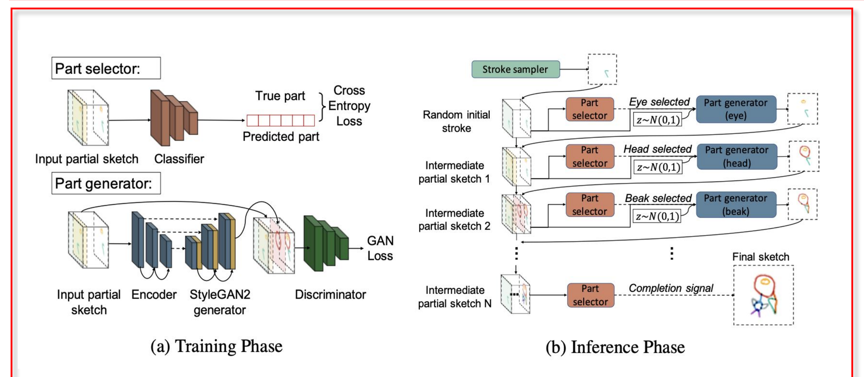
INPUT

- Sequence of Data Points

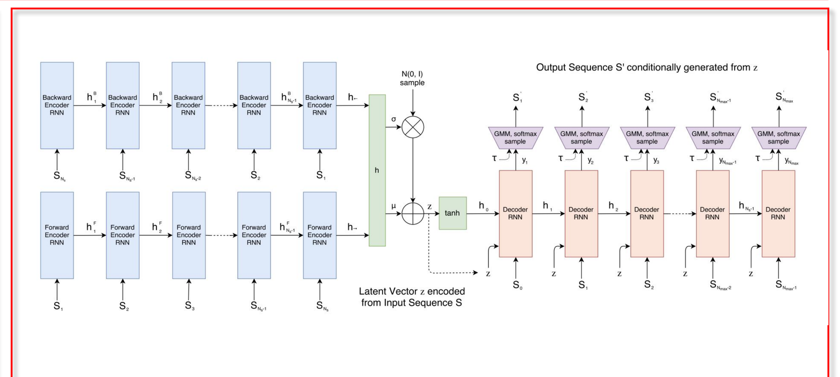
OUTPUT

- Model generated creative and novel sketches of the target category

LITERATURE REVIEW



DOODLERGAN: CREATIVE SKETCH GENERATION



A NEURAL REPRESENTATION of SKETCH DRAWING

DISCUSSION

Limitation:

Order of generated sequence: the output for sketch composer VAE, is a sequence of stroke-type-position data, where the order of the data matters. However, in practice, the order of the strokes is not important.

Spatial information: the data we use is vector sketch data, which means it is a sequence of points. This formulation neglects the spatial information of the sketch.

Future Improvement:

GNN: Graph Encoder/Decoder architecture can be used to achieve permutation invariance in stroke composition.

CNN: CNN of the 2D version sketch can be used to extract the spatial features.

ARCHITECTURE

StrokeRNN

Input:

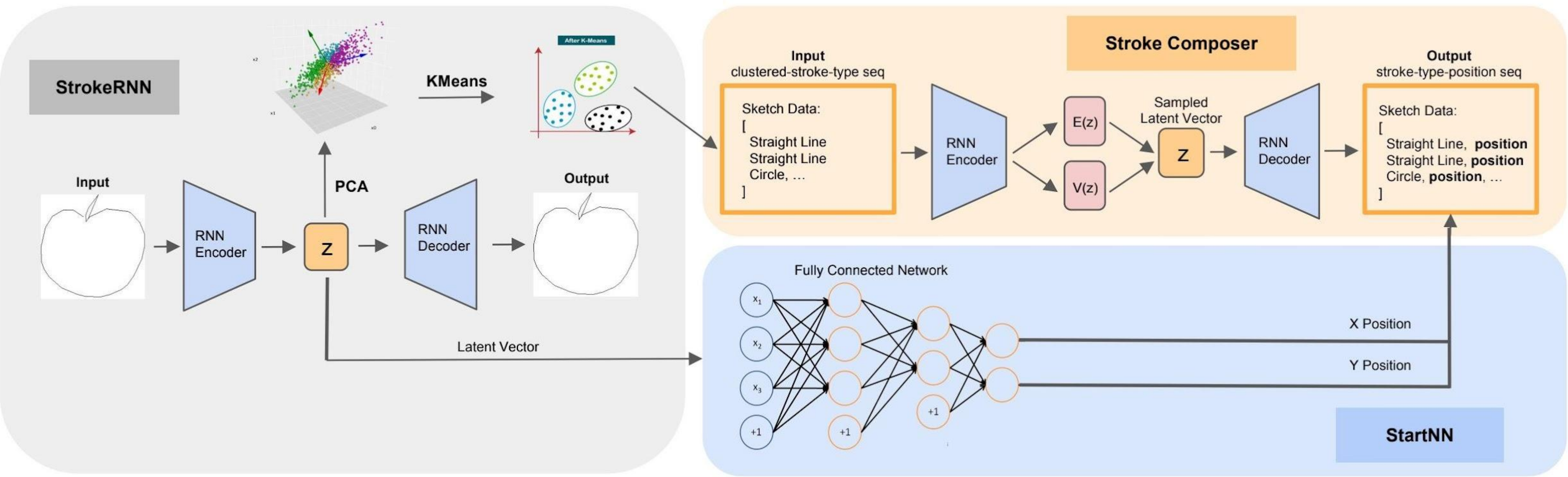
- stroke sequence data

Output:

- current stroke cluster id
- bottleneck latent vector

Loss Function:

$$L = \frac{1}{N}\sum_{j}^{N}\sum_{i}^{L}\|y_i - \hat{y}_i\|^2$$



Stroke Composer

Input: clustered-stroke-type sequence data

Output:generated stroke-type-position sequence

Loss Function:

$$l_i(\theta,\phi) = -\mathbb{E}_{z\sim q_\theta(z|x_i)}[\log p_\phi(x_i \mid z)] + \mathbb{KL}(q_\theta(z \mid x_i)\|p(z))$$

StartNN

Input: latent stroke vector from strokeRNN

Output: predicted start point absolute coordinates

Loss Funct

$$L = \frac{1}{N}\sum_{i}^{N}\|s_i - \hat{s}_i\|^2$$

RESULT

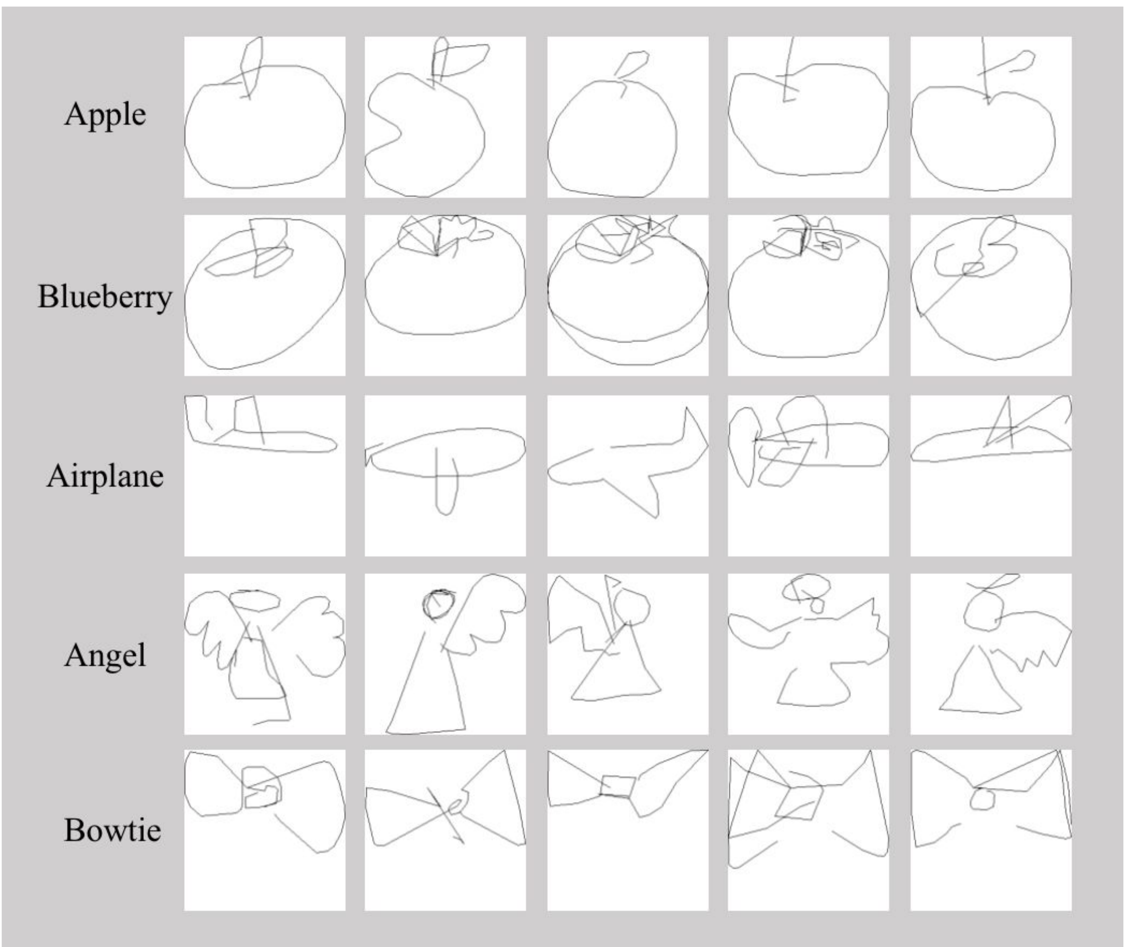


FIGURE 1. Sketch Composer Final Results for Different Category

	Bowtie	Angel	Airplane	Blueberry	Ant	Basket	Bed	Bird
Epoch 0	704.68	362.9	671.63	339.82	259.5	731.83	997.08	347.31
Epoch 500	37.64	13.63	37.9	15.75	**24.22**	34.22	64.38	13.85
Epoch 800	**34.25**	**13.45**	**36.71**	**15.73**	25.47	**30.85**	**47.05**	**11.38**

TABLE 1. StrokeRNN Reconstruct Loss for Different Categories on Different Epoch

	Bowtie	Angel	Airplane	Blueberry	Ant	Basket	Bed	Bird
Epoch 0	1.39	1.41	1.42	1.45	1.45	1.45	1.46	1.49
Epoch 500	**0.16**	0.24	**0.2**	0.24	0.22	**0.21**	0.22	0.23
Epoch 800	0.18	**0.22**	0.19	**0.22**	**0.21**	0.24	**0.20**	**0.23**

TABLE 2. Stroke Composer CE+KL Loss for Different Categories on Different Epoch

	Bowtie	Angel	Airplane	Blueberry	Ant	Basket	Bed	Bird
Epoch 0	15178.34	8490.41	21040.8	5250.58	40296.33	5981.5	8692.87	9356.87
Epoch 50	1595.27	1722.58	2219.32	4429.05	3170.22	1373.44	2565.85	2565.81
Epoch 100	**1340.86**	**1625.07**	**1691.57**	**3610.42**	**3147.34**	**1322.1**	**2116.45**	**2425.45**

TABLE 3. StartNN Mean Square Error Loss for Different Categories on Different Epoch

REFERENCE

[1] Ha, D., & Eck, D. (2017). A neural representation of sketch drawings. arXiv preprint arXiv:1704.03477.

[2] Ge, S., Goswami, V., Zitnick, C. L., & Parikh, D. (2020). Creative sketch generation. arXiv preprint arXiv:2011.10039.

[3] Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114.

[4] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

[5] Baggio, G., Van Lambalgen, M., & Hagoort, P. (2012). The processing consequences of compositionality. In The Oxford handbook of compositionality (pp. 655-672). Oxford University Press.

[6] J. Jongejan, H. Rowley, T. Kawashima, J. Kim, and N. Fox-Gieg. The Quick, Draw! - A.I. Experiment. https://quickdraw.withgoogle.com/, 2016.