Quick, Poem! 赋歌小画

Presented by

Ren Xuhua 任旭华 (SJTU, renxuhua@sjtu.edu.cn)
Liu Jingyuan 刘婧媛 (HKUST, jliucb@connect.ust.hk)
Hu Gansen 胡淦森 (SJTU, hugansen@sjtu.edu.cn)

Introduction

Visual arts convey rich emotional connotations, even by the simplest form of freehand drawings. In this project, we present a deep neural network that generates poems from freehand sketches.

Our Motivation

Our project is inspired by a previous work that generates poems from images [1]. We changed the input domain such that an end user can easily generate an interesting poem by drawing casual sketches of an object or a scene.

Methodology

Embedding Sketches and Poems

The designed network adopt InceptionNet to extract sketch representation and BERT to extract poem representation for building a paired embedding space.

Generating Poems

Poems are generated using LSTM and discriminated with respect to relevance with sketches and poeticness.

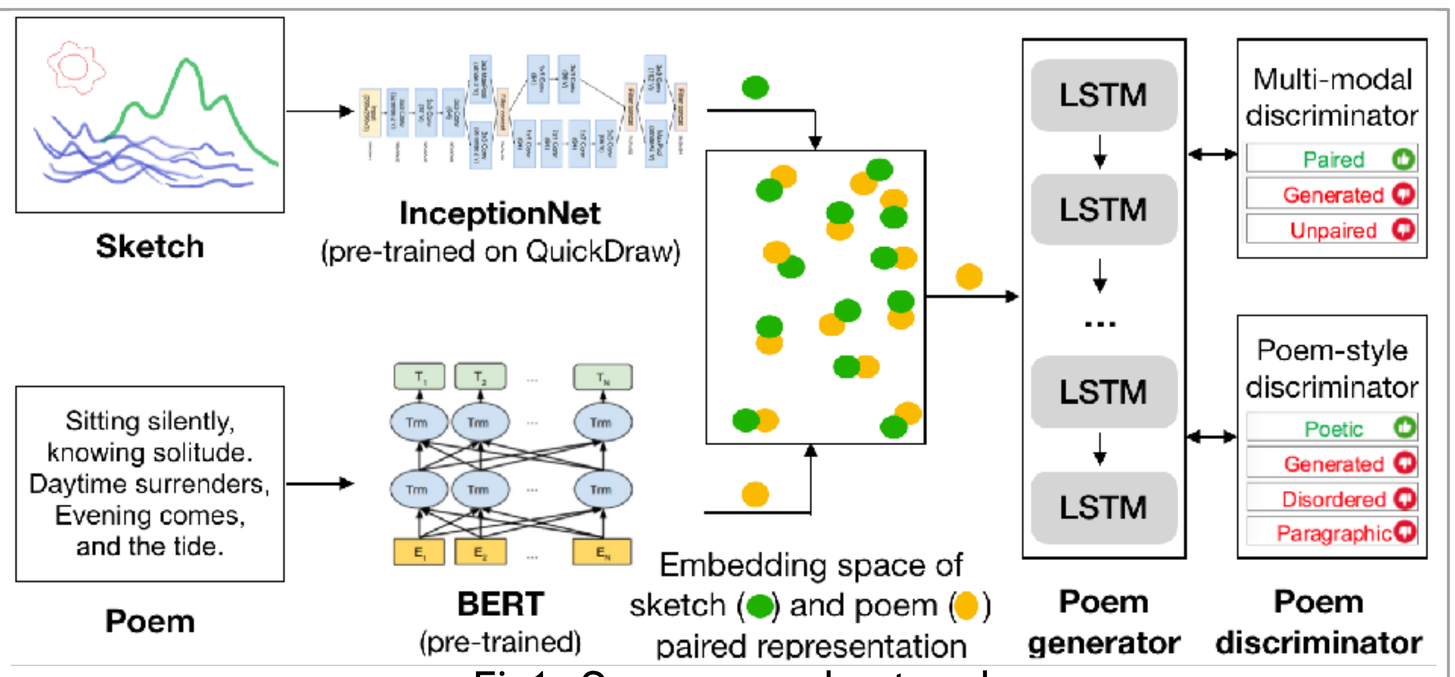


Fig1. Our proposed network.

Implementation

Challenges

1. There is no existing or similar dataset that contains paired sketch image and poems;
2. The designed network is too large to be trained within these few days.

Solution

We adopted an alternative by-passed network architecture using image as transition. A sketch is input into SketchyGAN [2] (re-trained on sketch-image dataset) to be upsampled to a realistic image. Then we adopted a trained Fast-RCNN [3] to visualize image contents. Finally inputs and labels are input into I2P-GAN [4] to generate poems.

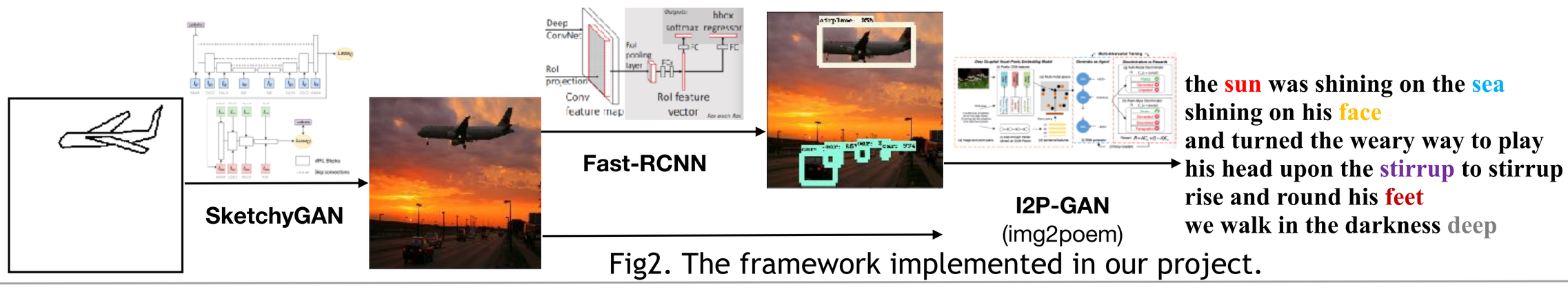
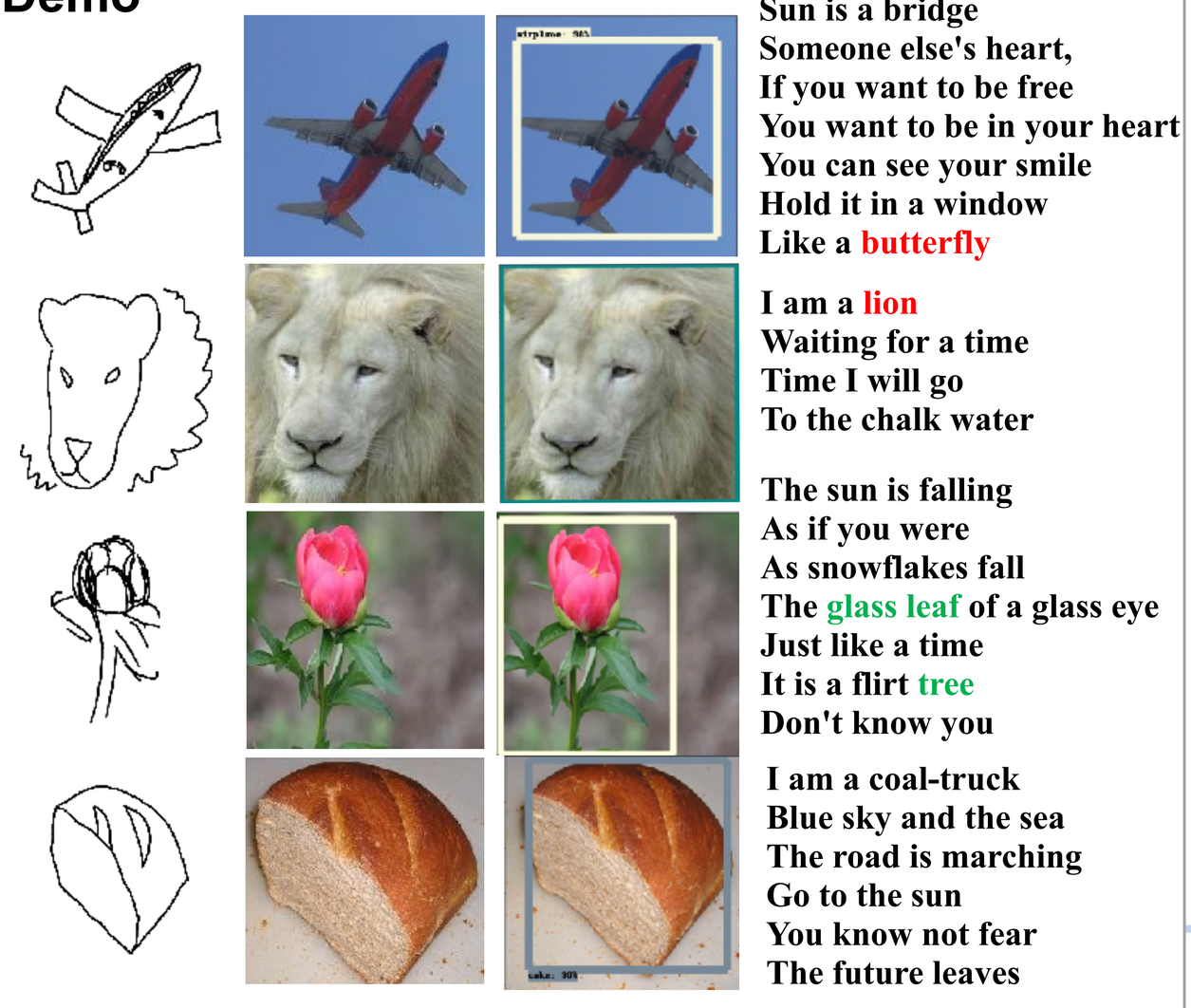


Fig2. The framework implemented in our project.

Demo



(a) sketch (b)image (c)poem clue (d) generated poem

Fig3. Demo gallery of *Quick, Poem!*

Future Work

1. The poetic clues detected by Fast-RCNN can potentially assists the poem generation process. I2P-GAN can be improved by augmenting with objects labels input.
2. Current available datasets contain sketches of single objects. By enabling multi-object sketch classification [5], the network supports generating poems from more complex scene sketch.

Reference

[1] Liu, B., Fu, J., Kato, M. P., & Yoshikawa, M. (2018). Beyond Narrative Description: Generating Poetry from Images by Multi-Adversarial Training. *arXiv preprint arXiv:1804.08473*.
[2] Chen, W., & Hays, J. (2018, January). SketchyGAN: Towards Diverse and Realistic Sketch to Image Synthesis. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 9416-9425).
[3] Girshick, R. (2015). Fast r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 1440-1448).
[4] Liu, B., Fu, J., Kato, M. P., & Yoshikawa, M. (2018). Beyond Narrative Description: Generating Poetry from Images by Multi-Adversarial Training. *arXiv preprint arXiv:1804.08473*.
[5] Zhang, J., Chen, Y., Li, L., Fu, H., & Tai, C. L. (2018, August). Context-based sketch classification. In Proceedings of the Joint Symposium on Computational Aesthetics and Sketch-Based Interfaces and Modeling and Non-Photorealistic Animation and Rendering (p. 3). ACM.