2 Related work

Generating poem is popular in recent years, and there are many people research on it. On the other hand, explain image with sentence has also made great progress after deep learning widely used. Our approach is a combination of above two aspects, we select two methods that performing well in corresponding aspects, and try to get a fantastic result.

3 Image to Sentences

The method we want to use here is proposed by Junhua Mao[1]. We use a multimodal Recurrent Neural Networks (m-RNN, [1]) model to address both the task of generating novel sentences descriptions for images, and the task of image and sentence retrieval. The whole m-RNN architecture contains a language model part, an image part and a multimodal part. The language model part learns the dense feature embedding for each word in the dictionary and stores the semantic temporal context in recurrent layers. The image part contains a deep Convulutional Neural Network (CNN) [3] which extracts image features.

4 Translate and Extract

Now we get the sentence, as to generate the meaningful poetry, we need to extract keywords from the sentences. We can use a translation process to deal the sentence and only use the words that appears in Chinese poetry dataset that we used in poem generator.

4 Poem generator

The method we will use to generate poem is proposed by Mirella Lapata[2].

First, to create the first line of poem, We select all phrases corresponding to the users’s keywords and generate all possible combinations satisfying the tonal pattern constraints. We use a language model to rank the generated candidates and select the bestranked one as the first line in the poem. In implementation, we employ a character-based recurrent neural network language model.

And after that, we generate the rest poem lines with original lines. Convert lines s1~si into vectors v1~vi with a convolutional sentence model(CSM). Next, a recurrent context model (RCM) takes v1~vi as input and outputs uji.Finally, u1i, u2i,...,ujI and the first j characters w1~wj in line S(i+1) serve as input to a recurrent generation model (RGM) which estimates P(w(j+1) = k| w(1~j) ,S(1~i) ) with k ∈V, the probability distribution of the (j+1)th character over all words in the vocabulary V.

5 Dataset

The dataset we used is same as the above two methods. First is IAPR TC-12 dataset, which used to training the model that explain image. The other is Chinese Poetry corpus(in Mirella Lapata, [2]).

[1] Junhua Mao, Wei Xu, Yi Yang, Jiang Wang, Alan L. Yuille : Explain Images with Multimodal Recurrent Neural Networks.

[2] Xingxing Zhang, Mirella Lapata : Chinese Poetry Generation with Recurrent Neural Networks.

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In NIPS, pages 1097–1105, 2012.