Summer and Evelyn Hasama
AIT-Budapest
May 2023

Haiku Generator From Photo

Introduction

This project aims to generate a haiku (a Japanese 5-7-5 syllable poem often relating to nature) using deep learning. To make things a bit more interesting, we are going to take a photo as input and use an existing object detection model to determine keywords. We then use these keywords in our haiku text generating model which we fine tuned on GPT-2 to output a haiku related to our input photo.

Previous Solutions

This is not the first time haiku generation has been done using deep learning principles. In 2018, software engineer and data scientist Jeremy Neiman, explored how deep learning can follow constraints, such as the fixed syllable structure, in Generating Haiku with Deep Learning. In Neiman's project, he uses an LSTM to build haikus from a single starting letter.

Additionally, in 2021, Soichiro Yokoyama from Hokkaido University created an AI system for Generating and Evaluating Japanese Haiku Poems with Deep Language Models. In his model, Yokoyama composes haikus using his GPT-2 model that is finetuned with a haiku corpus.

More recently, in 2022, Robert A. Gonsalves continues exploring haiku generation in Deep Haiku: Teaching GPT-J to Compose with Syllable Patterns. In his work, Gonsalves uses his fine tuned GPT-J 6B model to generate a haiku for a given topic using text and phonemes.

Dataset

We used the Haiku Dataset on Kaggle which contains 11,269 haikus scraped and cleaned from the haiku subReddit. We also used another Kaggle Haiku Dataset which contains over 118,000 English haikus gathered from multiple sources, only about 5,000 of which are used in training

the model. For our keywords, we have 60 nature related words generated from OpenAI's ChatGPT. The photos used later in the project to generate haikus from images are sourced by Google Images.

Proposed Method

First, we cleaned our haiku dataset and used a GPT-2 tokenizer which is based on byte-level Byte-Pair-Encoding to encode each of our poems. We then used the PyTorch Trainer to fine tune the GPT2LMHeadModel. Next, with this fine-tuned model, we generated haikus from our nature related words and filtered out poems that do not follow the 5-7-5 syllable structure using the syllables Python library.

Onto the application of our model, we used ImageAI's Image Classification to provide us with image recognized objects in each input photo. We used these objects as keywords for our haiku generator.

Evaluation Method

We evaluate the poems generated using BERTScore, described in the paper BERTScore: Evaluating Text Generation with BERT. The BERTScore has been shown to resemble human judgment on sentence-level and system-level evaluation. While the BERTScore computes precision, recall, and F1 measure, we evaluate the haikus solely on the mean of the F1 measure, as it combines both precision and recall. We compute the F1 measure for each generated haiku relative to 20 random reference haikus from our dataset.

Results

From all of the 60 nature keywords, we had 117 haikus generated. The average BERTScore was 0.55. Some example haikus include:

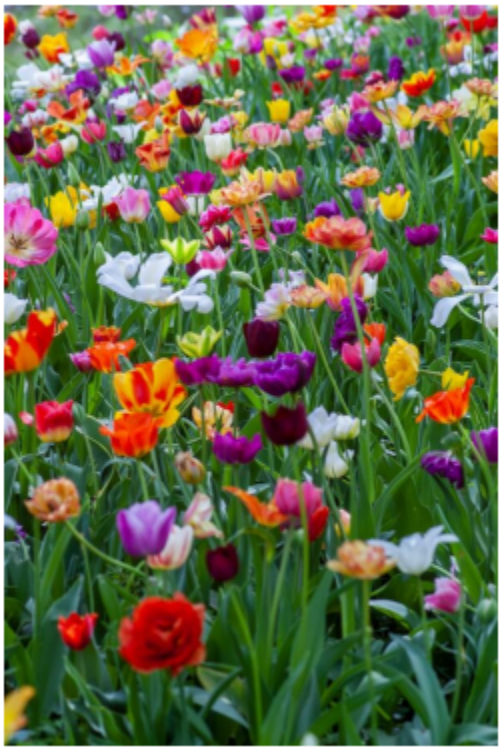
```
root in the garden
a small flower grows with me
to be a small bud

winded from the sun
```

```
the wind blows the earth silent
the sea is waiting

bright stars in the night
the light is fading away
to the sky just gone
```

When generating haikus from the nature images, we decided to select the top 2 keywords from each image. For each keyword, we display the poem with the highest BERTScore. Some examples include:



```
daisy in the tree
as the birds sing in the wind
my heart is the same
```

```
sandbar on the beach
the sandbar is too close
to the ocean floor
```

Discussion

Through this project we were able to generate haikus from photos. While they did sound like haikus, were often somewhat related to the photo, and followed the syllable pattern, unfortunately they did not make the most sense. If we were to continue with this project, the next step would be to further fine tune the model such that the haikus generated read more sensibly. We could also conduct a Mean Opinion Score (MOS) subjective test for more human evaluation.