

Computer Describes Pictures to You

Ruicheng Yuan
MSc in Big Data

Introduction

The aim of this dissertation project is to study image processing technology, deep learning convolution neural network and recurrent neural network, the inter-modal of computer vision and natural language process, and create an image automatic captioning algorithm. Use some experimental pictures to train and test the model.

Dataset

Flickr8K is a images with corresponding image captions dataset. It concludes 8000 images and every images have 5 different captions



A grey bird stands majestically on a beach while waves roll in .
A large bird stands in the water on the beach .
A tall bird is standing on the sand beside the ocean .
A water bird standing at the ocean 's edge .
A white crane stands tall as it looks out upon the ocean .

Applications



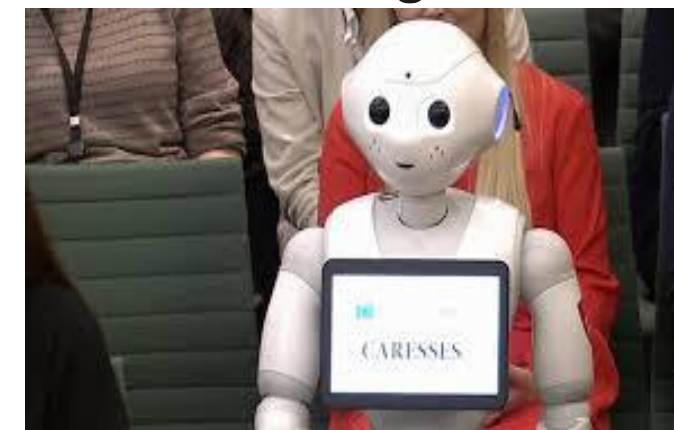
Drone Identifies Objects



Self driving car



Helping visually impaired



AI robots

Approach

Building model:

Both the CNN and RNN(LSTM) models were built using python 3.7 with Tensorflow and keras with GPU acceleration. using CNN to extract image features and LSTM to generate captions

Test model:

Using BLEU-n to test model's accurate.

