أكاديمية سدايا
SDAIA Academy

IMAGE CAPTIONS

PROJECTS PROPOSAL -T5 BOOTCAMP

Batoul Alosaimi
Amal Altamran
Amirah Alotaibi
Alanoud Alhussain
Shoroq Almutiri
Norah AlQahtani
Asma Alsulami

IMAGE CAPTIONS

Table of Contents

Introduction

Image caption Generator is a popular research area of Artificial Intelligence that deals with image understanding and a language description for that image. Automatic generation of an image description requires both computer vision and natural language processing techniques.
And from this point we decide to help people with disabilities as well as early childhood education By doing image caption English and Arabic and transformed to voice using artificial intelligence technology so, their lives can be improved.

Methodology

1. Using deep Learning we are planning to create a model that analyze images and describe them (create caption) in Arabic.
2. Create another model that analyze images and describe them (create caption) in English.
3. Convert written text into voice.

Dataset for English version

In the Flickr8k dataset, each image is associated with five different captions that describe the entities and events depicted in the image that were collected. By associating each image with multiple, independently produced sentences, the dataset captures some of the linguistic variety that can be used to describe the same image.
Our dataset structure is as follows:-
Flick8k/
Flick8k_Dataset/ :- contains the 8000 images
Flick8k_Text/
Flickr8k.token.txt:- contains the image id along with the 5 captions in the English language.

Dataset for Arabic version

- Translated copy from the English data set Flickr8k

Tools and Algorithms

- Keras
- TensorFlow
- Scikit-Learn
- Convolutional Neural Networks
- Pandas, Numpy, matplotlib, seaborn
- pydotplus
- googletrans
- gTTS