



Images Caption

DEEP LEARNING PROJECT

By:

Alanoud Alhussain

Batoul Alosaimi

Norah Alqahtani

Asma Alsulami

Shoroq Almutiri

Amal Altamran

Amirah Alotaibi

SDAIA Academy



Table of Contents

| | |
|----------------------------------|----------|
| <i>Introduction.....</i> | <i>2</i> |
| <i>Methodology.....</i> | <i>2</i> |
| <i>Dataset.....</i> | <i>2</i> |
| <i>Tools and Algorithm</i> | <i>2</i> |

Introduction

Image caption generation has gathered widespread interest in the artificial intelligence community. Automatic generation of an image description requires both computer vision and natural language processing techniques. While there has been advanced research in English caption generation, research on generating Arabic descriptions of an image is extremely limited. In our Project we plan to propose a deep learning model to help for Arab Society. Moreover, we plan to help people with disabilities as well as early childhood education 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 English.
2. Translate English caption into Arabic Languages using google translate API.
3. Convert written text into voice.

Dataset

- By translating a widely used English dataset; Flickr8k: it has 8092 images Each has 5 captions
Contains humans and animal
- From www.flickr.com
- <https://www.kaggle.com/ming666/flicker8k-dataset>

Tools and Algorithms

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