أكاديمية سدايا
SDAIA Academy

IMAGE CAPTIONS

IMAGE CAPTIONS

Deep Learning Project

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

Algorithms and Tools

- Keras
- TensorFlow
- PIL
- collections, random, re
- Convolutional Neural Networks
- Pandas, Numpy
- matplotlib, seaborn
- pydotplus
- googletrans
- gTTS

Introduction

Data Set

Methodology

Result

Conclusion

Future Work

Image Captioning (IC)

Objective:

- Help disabilities
- Facilitate education

Prediction Caption: a group of young people are standing in a blurry near a wall

Out[83]:

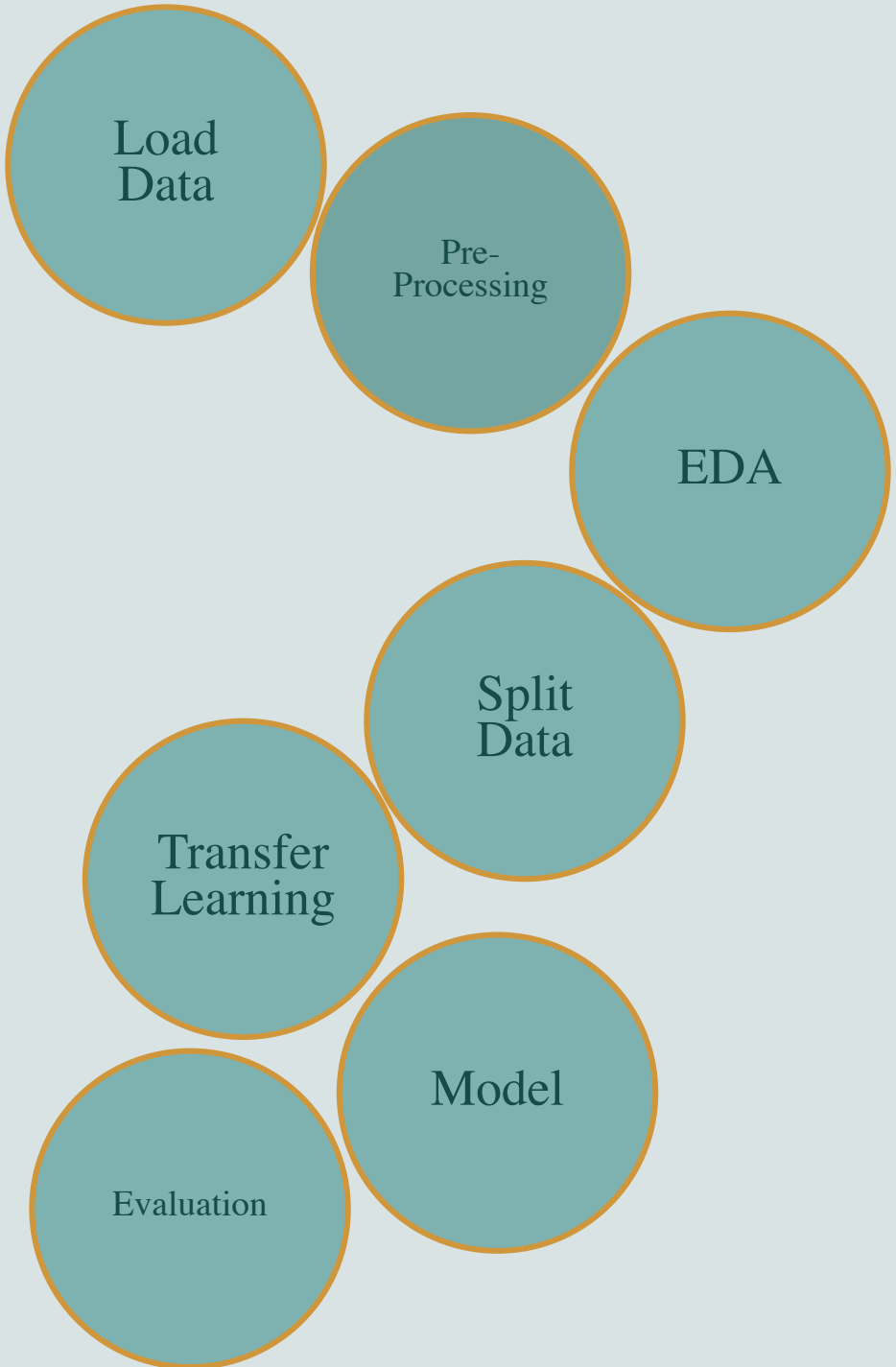
Our Class ☺

Dataset:

Flick8k_Dataset :- contains more than 8000 images

- Contains the image id along with the 5 captions in the English language.
- Contains the image id along with the 3 captions in the Arabic language.

Methodology



Exploratory Data Analysis

فتاة صغيرة مغطاة بالطلاء تجلس أمام قوس قزح
فتاة صغيرة تجلس أمام قوس قزح ملون كبير
فتاة أمام لوحة قوس قزح

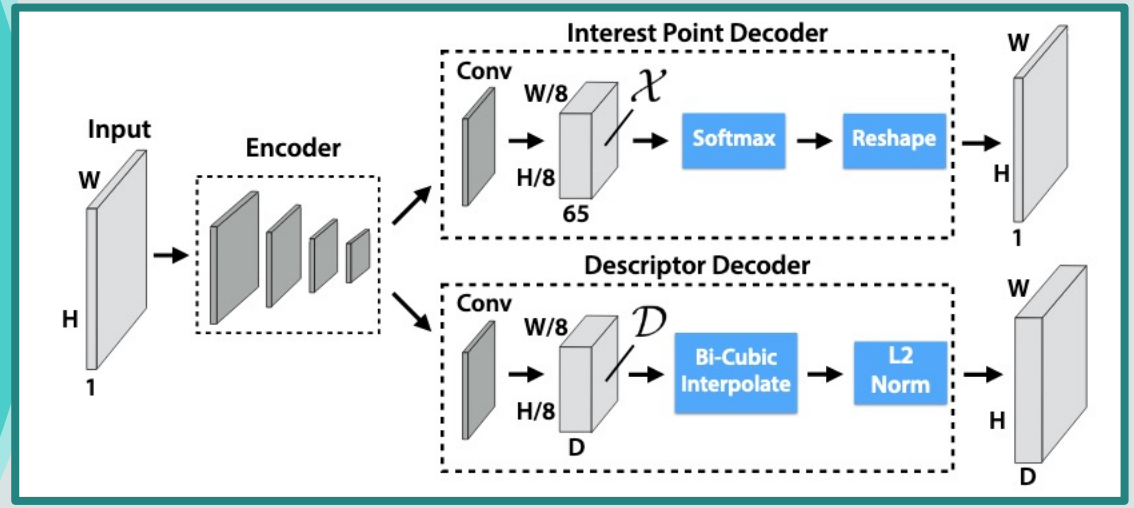
Figure 6: image from Arabic data set

A little girl covered in paint sits in front of a painted rainbow with her hands in a bowl .
A little girl is sitting in front of a large painted rainbow .
A small girl in the grass plays with fingerpaints in front of a white canvas with a rainbow on it .
There is a girl with pigtails sitting in front of a rainbow painting .
Young girl with pigtails painting outside in the grass .

Figure 7: image from English data set

Model used:
Transfer Learning

Inception V3: Is a Convolutional Neural Network for Assisting in Image Analysis and Object Detection.



Using four Experiment:

- Arabic : we tried two different Epochs 8, 15
- English : we tried two different Epochs 8, 15

Model Evaluation

BLEU – Score:
(BiLingual Evaluation Understudy) evaluating machine-translated text

Audio Libraries

!pip install googletrans
!pip install gTTs

gTTS Google Text to Speech, it is a Python library to interface with Google Translate text to speech API.

Note : This slide is attached to the audio

Result

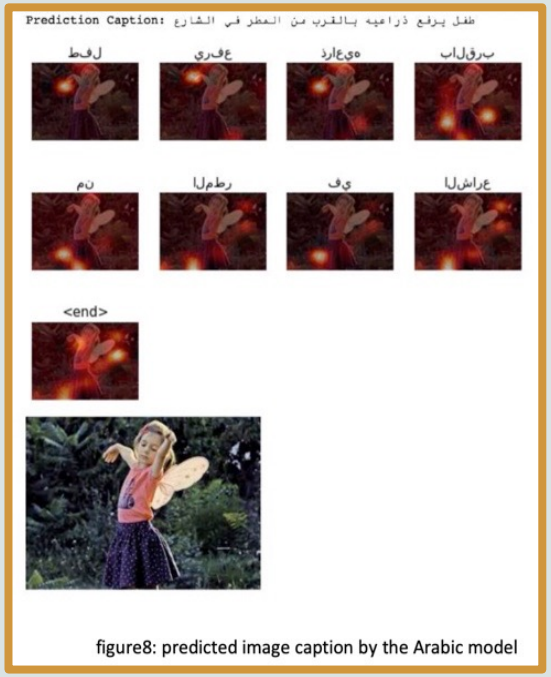


figure8: predicted image caption by the Arabic model

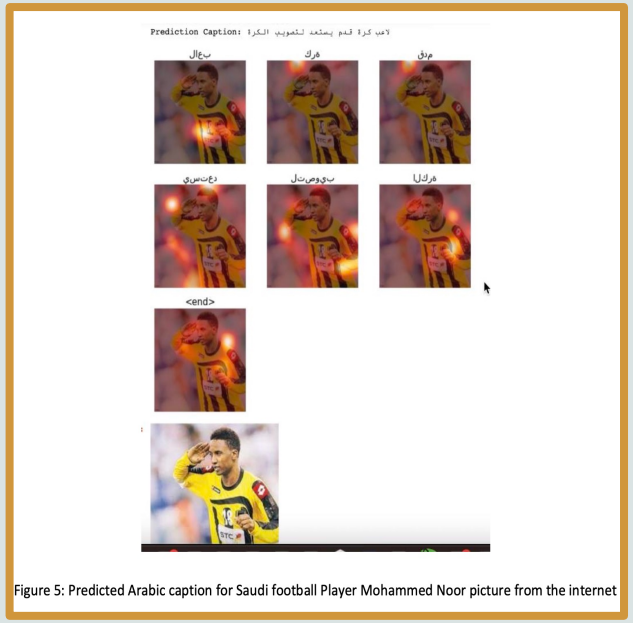


Figure 5: Predicted Arabic caption for Saudi football Player Mohammed Noor picture from the internet

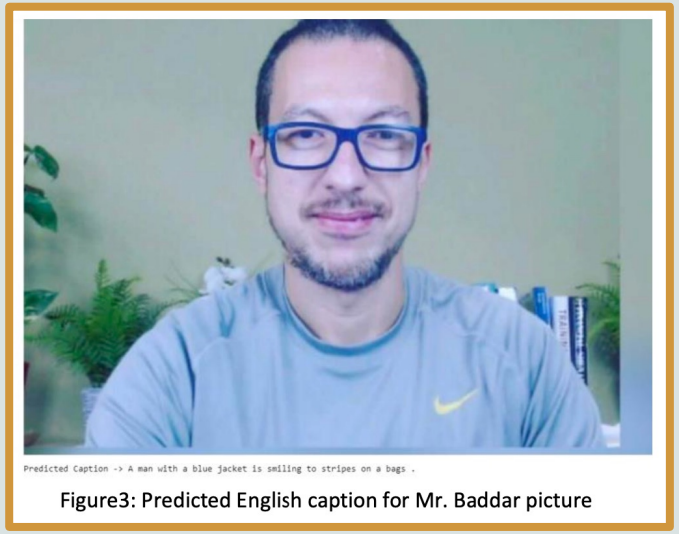


Figure3: Predicted English caption for Mr. Baddar picture

Figure: predict image caption by the English model

Conclusion

- We created a model that predict image description (caption) so people who cannot see will be able to know what inside images
- In addition, we improved a program that convert caption text into voice.
- Our project support both Arabic and English languages.

Future Work

Work on a larger dataset that contain local images
Improve Arabic voice model
Create website

Thank You

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

IMAGE CAPTIONS