IMAGE CAPTIONING USING CNN AND RNN

Members:

Akkem .Hema Bhargavi

Jagan Mohan Reddy Dwarampudi

Sowmiya TN

Ganesh G

Mentors: Naveen Nanda

Akshita Mehta

Introduction

- The purpose of this model is to generate captions for an image.
- Image captioning aims at generating captions of an image automatically using deep learning techniques such as CNN and RNN.

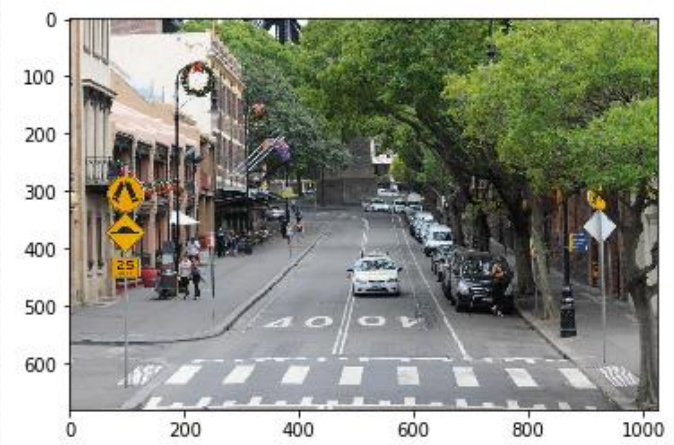
Dataset Used

- Traffic signs dataset from Open Images Dataset.
- The general images from Flickr 8k Dataset.

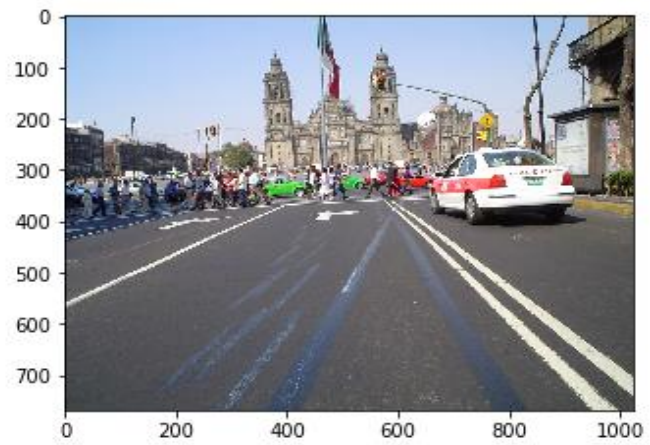
Methodology / Model Used

- In this project, two models : CNN and RNN are used.Initially, the objects in the image are detected using a **Convolutional Neural Network (InceptionV3).**
- Using the objects detected , a syntactically and semantically correct caption for the image is generated using **Recurrent Neural Networks (LSTM)** with attention mechanism**.**

Results Achieved



Greedy: speed ahead



Greedy: two people are walking past bus

Conclusion

- In our project, we have developed a model to caption the images. We have done research in order to understand our models in depth and have executed each model separately.
- We learned how the deep learning techniques work and how to create these models. We faced many challenges while running the model and with our datasets But later on we learned how to rectify the mistakes.

References

- Andrej Karpathy ,Li Fei-Fei,“DeepVisual-Semantic Alignments for Generating Image Description”, Department of Computer Science, Stanford University, published in IEEE Xplore 2015.
- Oriol Vinyals ,Alexander Toshev ,Samy Bengio,Dumitru Erhan ,“Show and Tell: A Neural Image Caption Generator”,arXiv 2015.
- KelvinXu ,JimmyLeiBa ,RyanKiros ,KyunghyunCho ,AaronCourville , RuslanSalakhutdinov, RichardS.Zemel Z,YoshuaBengio, “Show,AttendandTell:NeuralImageCaption GenerationwithVisualAttention”,arXiv 2015.
- Rahul Singha, Aayush Sharmaa, “Image captioning using Deep Neural Networks” ,Iowa State University,published in Research gate 2018.

Thank You