19CSE456
NEURAL NETWORKS AND DEEP LEARNING

IMAGE CAPTION GENERATOR

Name	Roll Number
Adarsh Thamaraiselvam	CB.EN.U4CSE19404
Matam Bhavya Sree	CB.EN.U4CSE19438
Mugunth J C	CB.EN.U4CSE19440
Nikhil Jeikar P M	CB.EN.U4CSE19443

INTRODUCTION

ABSTRACT

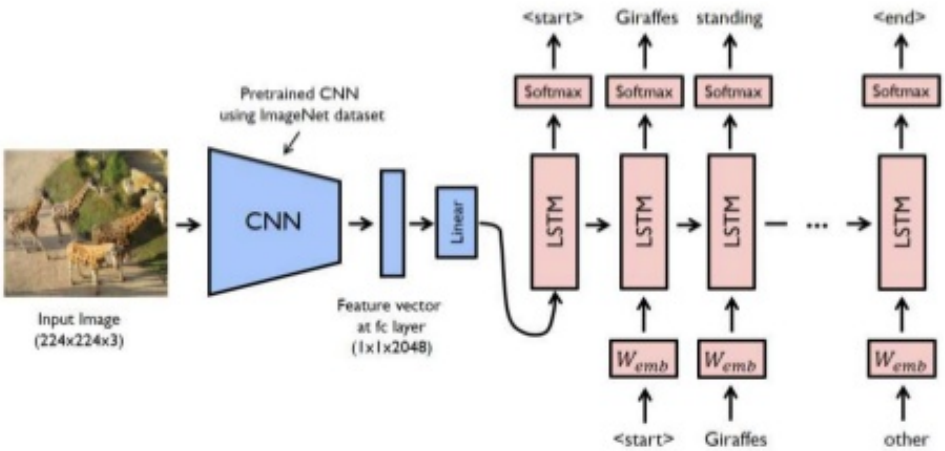
The human brain is capable of identifying and labelling an image when it is shown in front of us. With the advancement of Deep learning and Computer Vision, it is possible for a computer to do the same. Image Caption Generator is a process of identifying the context of an image and labelling it with appropriate caption. The image is labelled with English keywords using the datasets provided during model training.

Along with Deep Learning, it requires methods from both Computer Vision to comprehend the image's content and a language model from Natural Language Processing to convert the comprehension into words in the correct sequence. The model consists of Convolutional Neural Network as well as Recurrent Neural Network. The CNN is used for feature extraction from image and RNN is used for sentence generation.

We use the Flickr30k dataset for model training. It consists of 31.8k images and 158k captions giving different captions for the same image. This gives us a wide way to analysis and describe the image as output have wide choice it gives more data to work with.

Link : https://www.kaggle.com/hsankesara/flickr-image-dataset

TENTATIVE ARCHITECTURE



SOURCE::https://stats.stackexchange.com/questions/387596/image-caption-generator