IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 3, April 2023
Impact Factor: 7.301

Image Caption Generator using Deep Learning Approach

Prof. A. S. Narote[1], Kunal Vispute[2], Harshit Himanshu[3], Rajas Bhagatkar[4], Sneha Jadhav[5]
Professor, Department of Information Technology[1]
Students, Department of Information Technology[2,3,4,5]
Smt. Kashibai Navale College of Engineering Vadgaon Bk., Pune, Maharashtra, India

Abstract: *The creation of captions for a picture is the focus of the image caption generator. The image's semantic meaning is extracted and translated into plain language. Also, there are built-in programs that create and offer a caption for a certain image. Image captioning is the process of creating a description of a picture in the form of a caption. The system must recognize and develop connections between things, people, and animals. This study uses deep learning to find, identify, and produce interesting captions for a given image. The practice of creating textual descriptions of a given image using computer vision and natural language processing methods is known as image captioning. This suggests a strategy for deriving a caption for a picture that highlights particular objects in the image.*

Keywords: Deep Learning, Feature Extraction, thresholding, image segment, CNN model, NLP

I. INTRODUCTION

Image caption generation is a task that involves image processing and natural language processing concept to detect the context of an image and describe them in natural language.While human beings are able to do it easily,it takes a lot of computational power for a computer system. The mechanism must detect and establish relationships between objects, people, and animals. There are several steps to generating captions, such as understanding the visual representation of objects, establishing relationships among the objects and generating captions both linguistically and semantically correlated.

II. OBJECTIVES

- Our project aims to learn the concepts of CNN and LSTM models and build a working model of an Image caption generator by implementing CNN with LSTM.
- This System is based on pre-processing of an image caption generator which will generally result in less effort and time

III. RELATED WORK OR LITERATURE SURVEY

[1]" Image Caption Generator"
Author: Megha J Panicker and Vrinda Mathur
The CNN and the LSTM worked together in proper synchronization, they were able to find the relation between objects in images. Over the years several other neural network technologies have been used to create hybrid image caption generators, similar to the one proposed here

[2]"Comparative Evaluation of CNN Architectures for ImageCaption Generation"
Author: SulabhKatiyar and Samir Kumar Borgohain
Aided by recent advances in Deep Learning, Image Caption Generation has seen tremendous progress over the last few years. Most methods use transfer learning to extract visual information, in the form of image features, with the help of pre-trained Convolutional Neural Network models followed by transformation of the visual information using a Caption Generator module to generate the output sentences.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 3, April 2023
Impact Factor: 7.301

[3]Image Caption Generator
Author: Liya Sunny and Sara Susan Joseph
Image Caption Generation involves training a Machine Learning model to learn to automatically produce a single sentence description for an image. For human beings, it is a trivial task. However, for a Machine Learning method to be able to perform this task, it has to learn to extract all the relevant information contained in the image and then convert this visual information into a suitable representation of the image which can be used to generate a natural language sentence description of the image.

[4] Image Caption Generating Deep Learning Model
Author: Aishwarya Maroju, Sri Doma, Lahari Chandarlapati
Image captioning is the process of generating descriptions about what is going on in the image. By the helpof Image Captioning descriptions are built which explain about the images. Image Captioning is basically very much useful in many applications like analyzing large amounts of unlabeled images and finding hidden patterns for Machine Learning Applications for guiding Self driving cars and for building software that guides blind people. This Image Captioning can be done by using Deep Learning Models.

[5] Image Captioning using Deep Learning
Author:Murk Chohan, Adil Khan, Muhammad Saleem Mahar,Saif Hassan, Abdul Ghafoor, Mehmood Khan
Auto Image captioning is defined as the process of generating captions or textual descriptions for images based on the contents of the image. It is a machine learning task that involves both natural language processing (for text generation) and computer vision (for understanding image contents). Auto image captioning is a very recent and growing research problem nowadays. Various new methods are being introduced daily to achieve satisfactory results in this field.

[6] Bottom-Up and Top-Down Attention for Image Captioning and Visual Question Answering
Author: Peter Anderson, Xiaodong He, Chris Buehler, Damien Teney, Mark Johnson,Stephen Gould1 Lei Zhang
Top-down visual attention mechanisms have been used extensively in image captioning and visual question answering (VQA) to enable deeper image understanding through fine-grained analysis and even multiple steps of reasoning. In this work, we propose a combined bottom-up and topdown attention mechanism that enables attention to be calculated at the level of objects and other salient image regions.

[7] Image Caption Generator
Author: Parth Kotak , Prem Kotak
Automatically creating the description or caption of an image using any natural language sentences is a very challenging task. It requires both methods from computer vision to understand the content of the image and a language model from the field of natural language processing to turn the understanding of the image into words in the right order

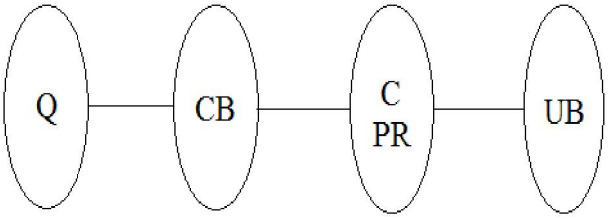
[8] Visual Image Caption Generator Using Deep Learning
Author: Priyanka Kalena,Nishi Malde ,Aromal Nair ,Saurabh Parkar
Image Caption Generation has always been a study of great interest to the researchers in the Artificial Intelligence department. Being able to program a machine to accurately describe an image or an environment like an average human has major applications in the field of robotic vision, business and many more. This has been a challenging task in the field of artificial intelligence throughout the years.

[9] Image Caption Generator Using Convolutional Neural Network Algorithm
Author: Shaik Parvez
It is a very difficult challenge to automatically describe an image using a sentence from any natural language, such as English. It necessitates knowledge of both natural language processing and picture processing. The fusion of computer vision and natural language processing has received a lot of interest recently thanks to the advent of deep learning. This field is exemplified by image captioning, which teaches a computer to understand an image's visual information using

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 3, April 2023
Impact Factor: 7.301

one or more phrases. In addition to the ability to recognize the item and the scene, high-level image semantics also needs the ability to analyze the state, the properties, and the relationship between these things.

IV. MATHEMATICAL MODELING



Where,

Q = dataset
CB = Preprocess
C = apply cnn algorithm
PR = Evaluation
UB = predict outcome

B] Set Theory

1) Let S be as system video stream i.e. URL.

S = {In, P, Op, Φ}

2) Identify Input In preprocess

In = {Q}

Where, Q = read the dataset

3) Identify Process P as feature extraction

P = {CB, C, PR}

Where, CB = System check and preprocess dataset

C = apply CNN

PR = Preprocess request

4) Identify predict output

Op = {UB}

Where, UB =predict outcome

After preprocessing the request, the system decides whether or not a particular link is education-related. If it is not educationally related then the system adds that link tothe blacklist.

Φ = Failures and Success conditions.

Failures:

1. Huge database can lead to more time-consuming to get the information.
2. Hardware failure.
3. Software failure.

Success:

1. Search the required information available in Datasets.
2. User gets results very fast according to their needs.

Space Complexity:

The space complexity depends on the Presentation and visualization of discovered patterns. The more the storage of data more is the more space complexity.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 3, April 2023
Impact Factor: 7.301

Time Complexity:

Check No. of patterns available in the datasets= n

If (n>1) then retrieving information can be time-consuming. So the time complexity of this algorithm is O(n^n)..

V. EXISTING SYSTEM AND DISADVANTAGES

There are some drawbacks to the present image captioning models. First, encoder–decoder models cannot distinguish between more important and less important objects in a given image since the caption is generated using only a feature vector for the entire input image.

Disadvantage:

- It required an internet connection must.

VI. ADVANCED SYSTEM AND ADVANTAGES

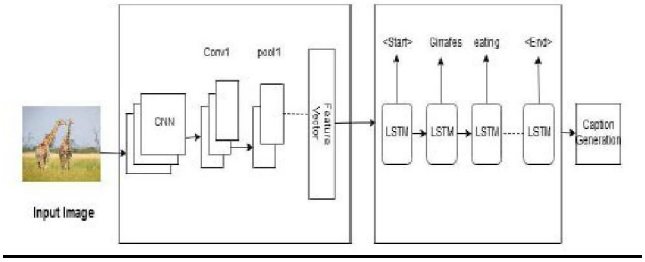


Figure: Advance System Architecture

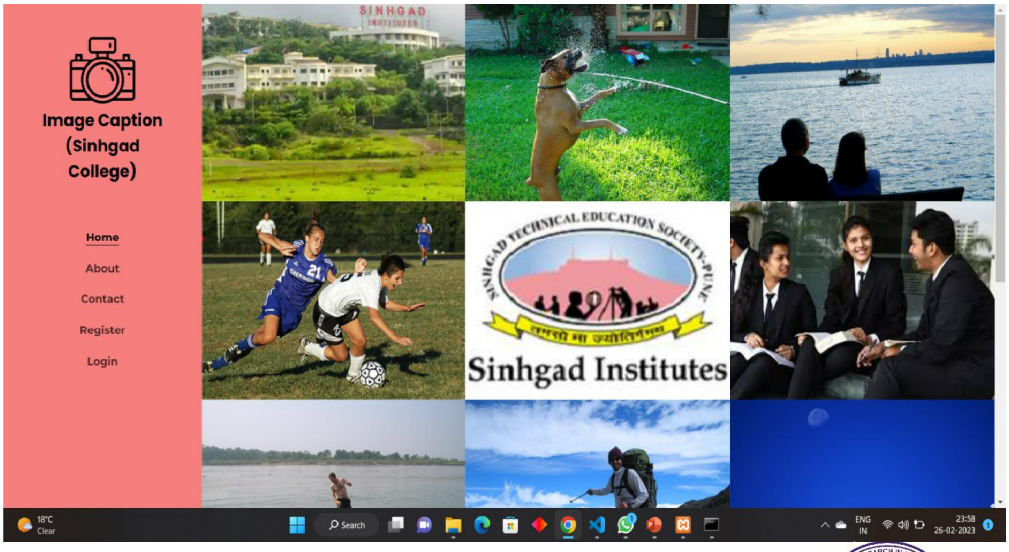
Advantages:

- In the qualitative results of our experiments, we demonstrated the advantages of our model in terms of generating descriptive captions and providing explanations for the output
- Secure and efficient system

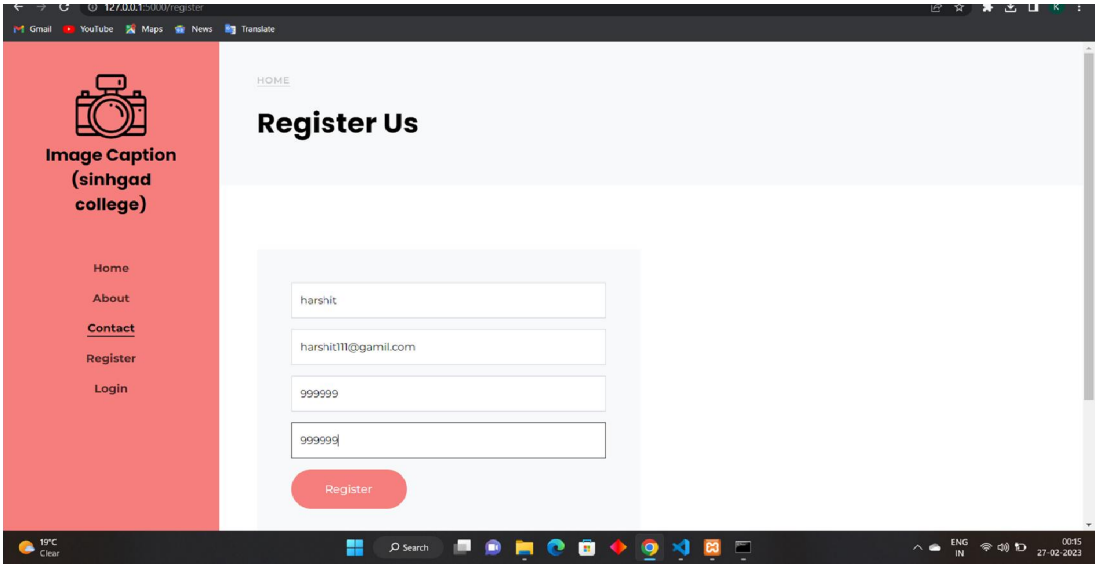
Found Result and Expected Result:

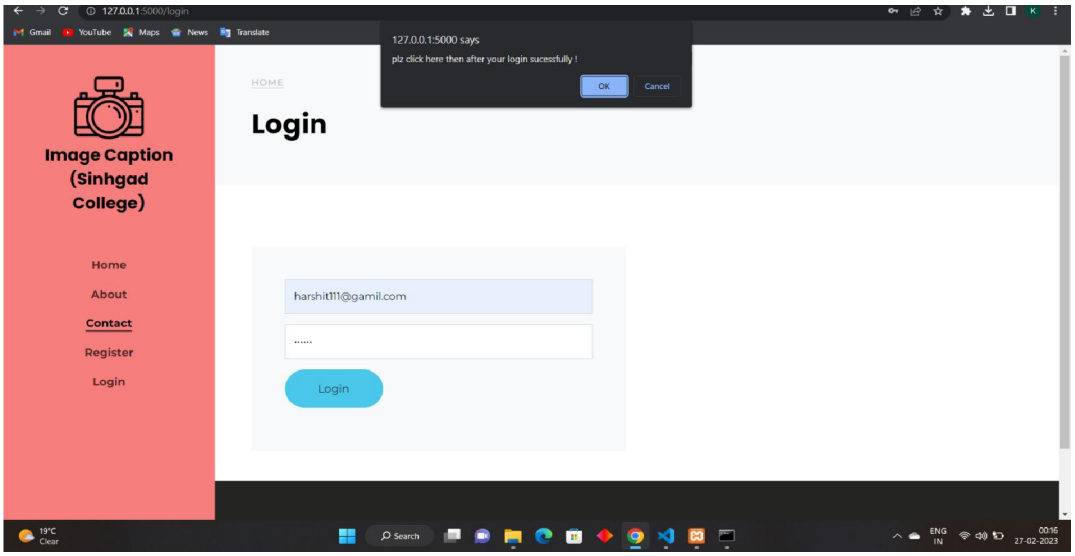
Our final model is efficient but it's not fast enough as we expected the model to retrieve the name of the image faster what we expected was a more efficient model

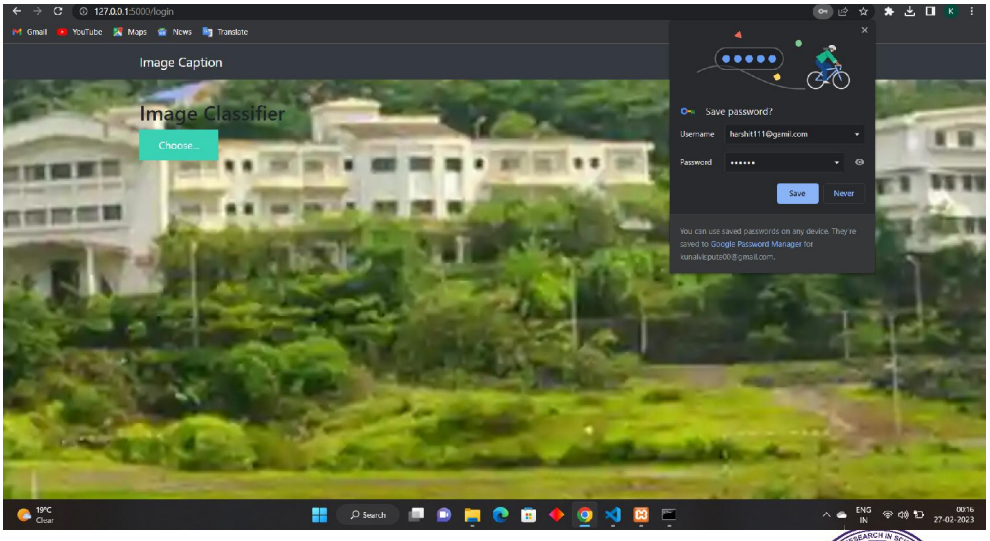
VII. RESULTS



IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 3, April 2023
Impact Factor: 7.301

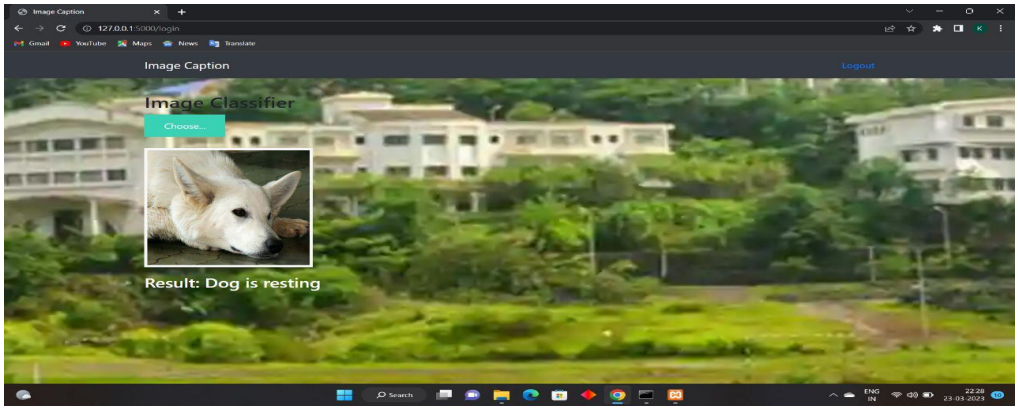






DOI: 10.48175/IJARSCT-9205

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 3, April 2023
Impact Factor: 7.301



VIII. CONCLUSION

This study offers a deep learning approach for creating image captions with neural networks; the suggested technique now incorporates a Flickr 8k dataset. Compared to existing image caption generators, the suggested deep learning technology produced captions with more descriptive meaning. Further research could lead to the creation of a hybrid photo caption generator model for more accurate captions.

REFERENCES

[1]. Megha J Panicker, Vikas Upadhayay, Gunjan Sethi, Vrinda Mathur: "Image Caption Generator" International Journal of Innovative and Exploring Engineering (IJITEE) JAN-2021.

[2]. Murk Chohan, Adil Khan, Muhammad Saleem Mahar, Saif Hassan, Abdul Ghafoor, Mehmood Khan: "Image Captioning using Deep Learning: A Systematic Literature Review" IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 11, No. 5, 2020

[3]. Peter Anderson, Xiaodong He, Chris Buehler, Damien Teney, Mark Johnson, Stephen Gould, Lei Zhang : "Bottom-Up and Top-Down Attention for Image Captioning and Visual Question Answering" 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition

[4]. Liya Ann Sunny, Sara Susan Joesph, Sonu Sarah Geogy, K.s>Ssreelakshmi, Abin T. Abraham: "Image Caption Generator" International Journal of Recent Advances in Multidisciplinary Topics (IJRSMT) APRIL-2021

[5]. Aishwarya Maroju, Sneha Sri Doma, LahariChandarlapati :"Image Caption Generating Deep Learning Model"International Journal of Engineering Research & Technology (IJERT), September-2021

[6]. Sulabhkatiyar,Samirkumarborgohain:"Comparative Evaluation of CNN Architectures for ImageCaption Generation"(IJACSA) International Journal of Advanced Computer Science and Applications,2020

[7]. Parth Kotak and Prem Kotak:"Image Caption Generator"(IJERT)International Journal of Engineering Research & Technology,November-2021

[8]. Priyanka Kalena ,NishiMalde ,Aromal Nair ,Saurabh Parkar:" Visual Image Caption Generator Using Deep Learning" International Conference on Advances in Science & Technology (ICAST-2019)

[9]. Shaik Parvez:"Image Caption Generator Using Convolutional Neural Network Algorithm" International Journal of Science and Research (IJSR)2022