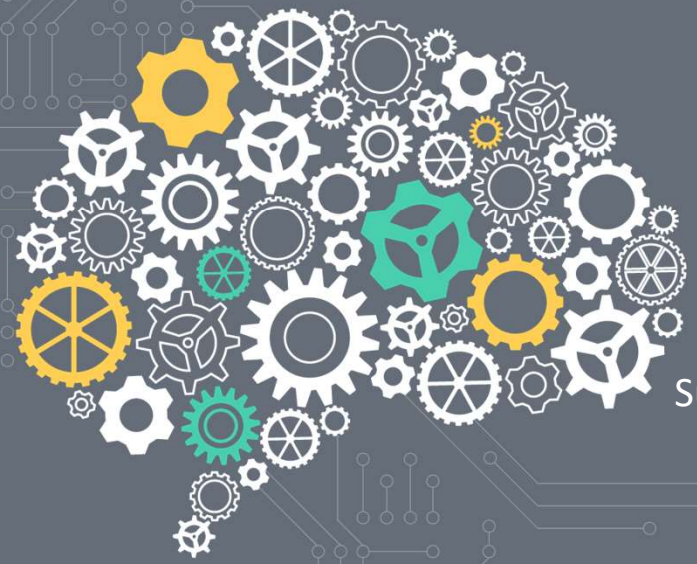
Welcome to the CSE-465 Project Presentation

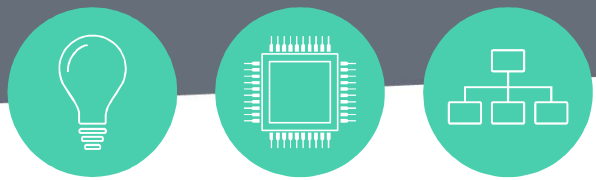


Project Title: Automatic Image Caption using CNN,LSTM ,VGG-16,VGG-19,ResNet-50,Inception-V3,OFA,ViT, BLIP.

Course: CSE-465

Section: 02

Submitted to: Dr. Mohammad Ashrafuzzaman Khan(Azk)



Name	Student ID
S M Gazzali Arafat Nishan	1831513642
Rofiqul Alam Shehab	1831185042
Shafiul Bashar	1831446642

What's the problem?

Image Caption from an image

Image Captioning is the process of generating a textual description for given images. And, It uses both Natural Language Processing and Computer Vision to generate the captions.

How did we solve the problem?

01

Dataset (Input Image)

We have used Flickr8K dataset for training images(8091 images).

02

Performed Data Cleaning

we have done data Cleaning by removing punctuations, words containing numbers and loaded dataset.

03

Extracted Feature vector from the images

Basically, we have pre_process the input images and extracted feature vector.

04

Train the model

Then, we have loaded the dataset for training model like CNN,LSTM,VGG-16,ResNet-50,Inception-V3

05

Tokenizing the vocubulary

Tokenizer class vectorised text corpus & each integer will represent token in dictionary.

06

Creating Data generator

Data generator is used bymodel.fitgenerator() ,retrieve photo features, encode the sequence.

07

Defining CNN-RNN model

Then, Features from the CNN model squeezed from 2048 to 256 nodes , LSTM sequence model, Merging both models.

08

TRAINING SECTION

We have trained our model by running (100 epochs) and generated our model.

09

Test The Model

We tested our model to evalute the captions.

Model Parameters

Dataset	Algorithms Used	Num of Batch Size	Num of Hidden Layers	Activation Function
Flickr 8k	1. CNN	256	3	ReLU
Flickr 8k	2. LSTM	256	3	SoftMax
Flickr 8k	3.ResNet-50	256	5	ReLU
Flickr 8k	4.VGG-16	256	0	SoftMax
Flickr 8k	5.VGG-19	256	0	SoftMax
Flickr 8k	6.Inception-V3	64	4	-
Flickr 8k	7.Custom Transformers	-	8	-

Result Analysis

Group Members	Datasets	Algorithms Used	Num of Epochs	Progress
Shafiul Bashar	Flickr 8k	1. CNN	10	Done
	Flickr 8k	2. LSTM	10	Done
Rofiqul Alam Shehab	Flickr 8k	3.ResNet-50	10	Done
	Flickr 8k	4.VGG-16	10	Done
	Flickr 8k	5.VGG-19	10	Done
S M Gazzali Arafat Nishan	Flickr 8k	6.Inception-V3	10	Done
	Flickr 8k	7.Custom Transformers	10	Done
	MS COCO	8.OFA	Pre Trained Model	Done
	Still in progress	9. Vit, Blip_Vit,	Pre Trained Model	

Image Caption from CNN-LSTM

```
from PIL import Image
img = Image.open('/content/drive/MyDrive/CSE-465/Flicker8k_Dataset/3275704430_a75828048f.jpg')
img
```

```
[ ] !python3 '/content/drive/MyDrive/CSE-465/testing_caption_generator.py' -i '/content/drive/MyDrive/CSE-465/Flicker8k_Dataset/3275704430_a75828048f.jpg'
    # !python3 '/content/drive/MyDrive/ML/testing_caption_generator.py' -i '/content/drive/MyDrive/ML/Flicker8k_Dataset/3738685861_8dfff28760.jpg'
```

```
2022-04-17 18:20:41.547727: W tensorflow/core/common_runtime/gpu/gpu_bfc_allocator.cc:39] Overriding allow_growth setting because the TF_FORCE_GPU_ALLOW_GROWTH environment variable i
2022-04-17 18:20:46.616605: W tensorflow/core/common_runtime/bfc_allocator.cc:343] Garbage collection: deallocate free memory regions (i.e., allocations) so that we can re-allocate a

start man in black shirt and black hat is sitting on the sidewalk end
```

Image Caption from VGG-16

```
[ ] from PIL import Image
    img = Image.open('/content/drive/MyDrive/CSE-465/Flicker8k_Dataset/111537222_07e56d5a30.jpg')
    img
```

```
[ ]
```

```
[ ] !python3 '/content/drive/MyDrive/CSE-465/testing_caption_generator.py' -i '/content/drive/MyDrive/CSE-465/Flicker8k_Dataset/111537222_07e56d5a30.jpg'
    # !python3 '/content/drive/MyDrive/ML/testing_caption_generator.py' -i '/content/drive/MyDrive/ML/Flicker8k_Dataset/3738685861_8dfff28760.jpg'
```

```
2022-04-17 17:47:38.611675: W tensorflow/core/common_runtime/gpu/gpu_bfc_allocator.cc:39] Overriding allow_growth setting because the TF_FORCE_GPU_ALLOW_GROWTH environ
2022-04-17 17:47:45.329493: W tensorflow/core/common_runtime/bfc_allocator.cc:343] Garbage collection: deallocate free memory regions (i.e., allocations) so that we ca


start man is climbing up rock end
```

Image Caption From Inception-V3 & Transformers

InceptionV3_Transfomer_Image_Caption_Transformer.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

```
            caption.remove(i)

    for i in real_caption:
        if i=="<unk>":
            real_caption.remove(i)

    #remove <end> from result
    result_join = ' '.join(caption)
    result_final = result_join.rsplit(' ', 1)[0]

    real_appn = []
    real_appn.append(real_caption.split())
    reference = real_appn
    candidate = caption
    score = sentence_bleu(reference, candidate, weights=(1.0,0,0,0))
    print(f"BLEU-1 score: {score*100}")
    score = sentence_bleu(reference, candidate, weights=(0.5,0.5,0,0))
    print(f"BLEU-2 score: {score*100}")
    score = sentence_bleu(reference, candidate, weights=(0.3,0.3,0.3,0))
    print(f"BLEU-3 score: {score*100}")
    score = sentence_bleu(reference, candidate, weights=(0.25,0.25,0.25,0.25))
    print(f"BLEU-4 score: {score*100}")

    print ('Real Caption:', real_caption)
    print ('Predicted Caption:', ' '.join(caption))
    temp_image = np.array(Image.open(image))
    plt.imshow(temp_image)
```

```
BLEU-1 score: 41.66666666666667
BLEU-2 score: 64.54972243679028
BLEU-3 score: 76.90182164956363
BLEU-4 score: 80.34284189446518
Real Caption: the old man has long white beard and yellow and black hat
Predicted Caption: man in black shirt and white shirt is standing on the street
<matplotlib.image.AxesImage at 0x7f273e773750>
```

Thank You